VOLUME 23 DECEMBER, 1959 NUMBER 4

An Investigation of ESP in the First and Second Grades of Dutch Schools.....*J. G. van Busschbach* 227

Exploring the Parameters of Research Variables
Gertrude R. Schmeidler 238

Some Problems in Group Testing of ESP
C. J. Adcock and D. Quartermain 251

Influence of the Position of the Subject in Relation to the Test Cards upon ESP Results.....*Rémy Chauvin* 257

Comments and Letters
Correspondence (*A. A. Mullin, T. Dumont, G. Zorab*) 267

Book Reviews

The Sixth Sense by Rosalind Heywood
Edmond P. Gibson 273

Water Witching U.S.A. by Evon Z. Vogt and Ray Hyman
Horace C. Levinson 274

Parapsychologie by G. Zorab.....*J. G. van Busschbach* 277

Parapsychological Abstracts..... 280

Announcements by the Parapsychological Association
The Second Annual Convention..... 282

Glossary 296

IENTIFIC QUARTERLY DEALING WITH EXTRASENSORY PERCEPTION, THE PSYCHOKINETIC EFFECT, AND RELATED TOPICS

DUKE UNIVERSITY PRESS
DURHAM, N. C.

The Journal *of* Parapsychology

EDITORIAL STAFF

J. G. PRATT, L. E. RHINE, and W. G. ROLL, *Editors*
DOROTHY H. POPE, *Managing Editor*
J. A. GREENWOOD and T. N. E. GREVILLE, *Statistical Editors*

This JOURNAL, published on the thirtieth day of March, June, September, and December, is devoted primarily to the original publication of experimental results and other research findings. In addition, articles presenting reviews of the literature relevant to parapsychology, criticisms of published work, proposals for research, theoretical and philosophical discussions, reports of spontaneous cases, and new methods of mathematical analysis will be published as JOURNAL space allows. The editors welcome comments, inquiries, and suggestions about the JOURNAL and any of its contents.

Contributions submitted for publication and all editorial communications should be addressed to the Managing Editor. Correspondence with the editors is advised before submitting articles other than reports of experimentation. Since it does not forward manuscripts by registered mail, the JOURNAL cannot guarantee that they will not be lost in transit. Contributors are therefore urged to keep copies of their papers. All manuscripts should be typewritten, double spaced. References should be given in the form adopted by the JOURNAL. Contributions will be acknowledged immediately upon receipt, and the editors will communicate further regarding each manuscript as soon as possible, generally within two months. Articles will be published in the order of final acceptance insofar as content balance in the JOURNAL will allow.

Reprints may be ordered when the proof is returned. The JOURNAL will bear one-half the cost of reprints up to two hundred copies.

The subscription price is $6.00 a year; single current numbers can be obtained for $1.50 each. The rate for back volumes is $7.00; for single back numbers, $1.75. All remittances should be made payable to the JOURNAL OF PARAPSYCHOLOGY. Missing numbers will be supplied free when lost in the mails if written notice is given within one month of the date of issue. Correspondence concerning subscriptions, change of address, back numbers, and other business communications should be addressed to the Managing Editor, Dorothy H. Pope, College Station, Durham, N. C.

Entered as second class matter May 21, 1942, at the post office at Durham, N. C., under the act of March 3, 1879.

DUKE UNIVERSITY PRESS

The Journal *of*
Parapsychology

VOLUME 23 DECEMBER, 1959 NUMBER 4

AN INVESTIGATION OF ESP IN THE FIRST AND SECOND GRADES OF DUTCH SCHOOLS

By J. G. VAN BUSSCHBACH[1]

ABSTRACT: Previous ESP tests carried out in public schools with children in the fifth and sixth grades gave significant results, while tests with more advanced pupils were at the chance level. This led the author to extend his experiments to younger children as a further part of his study of a possible parapsychical relationship between teacher and pupil. He carried out a total of 51,624 trials in the present tests with first- and second-grade subjects in the cities of Amsterdam and Dordrecht in Holland. The children were asked to guess the targets which their teacher was looking at as she sat concealed from their view. Each trial required a choice of one of three possible targets, and three types of targets were used in each class test. In one series of 12 trials the three targets were geometrical figures; in another, ordinary objects; and in a third, colors. The subjects indicated their selections by a "crossing-out" process.

There were statistically significant findings in three different aspects of the experiment. (1) With regard to the over-all results: the hits scored by the subjects in Amsterdam were not extrachance, but those in Dordrecht were extremely significant. The difference between the two, and the combination of the two were also highly significant. (2) With regard to grade differences: the first-grade subjects scored higher than the second-graders to a statistically significant degree. (3) With regard to differences between the sexes: the girls scored higher than the boys in both cities, and the difference between the totals for girls and boys in the combined experiment was statistically significant.

The author suggests that the higher scoring in Dordrecht could be due to a difference in the test procedures or to other factors, such as the manner in which the tests were presented by the test leaders in the two cities.—Ed.

INTRODUCTION

SINCE 1951 I have been investigating the question whether an extrasensory factor may be playing a part in the relation between educator and pupil. These investigations, as previously presented

[1] I wish to express my appreciation to the Foundation for the Study of Psycho-physics in the Netherlands for the generous contribution made toward the financing of the two investigations reported here. This report is based on a paper presented at the Parapsychological Association convention in Sept., 1958. The research was originally reported in the *Tijdschrift voor Parapsychologie*, 1958, No. 4-5, 172-78.

in three reports in this *Journal* (**4**, **5**, **6**), have included a total of 131,500 GESP trials in city schools in different places in the Netherlands and in North Carolina. This work was made possible through the co-operation of 148 teachers, and a total of 4,396 pupils took part as subjects. The tests that were made with pupils of the fifth and sixth grades of city schools with the classroom teacher acting as the agent in the experiment (consisting of 83,230 trials from the total indicated above) gave a highly significant result (P=.000006), whereas the remaining trials made under other conditions (higher classes, secondary schools, or with either a pupil or a stranger acting as agent) came out with insignificant results.

These results stimulated me to extend the experiment also to younger children, and for this purpose I chose a number of first and second grades with pupils of 6 to 8 years in various schools in Amsterdam (1956) and Dordrecht (1957). The present paper is a report on this additional study of the parapsychical factor in the teacher-pupil relationship.

Procedure

The carrying-out of the research plan involved special difficulties, many of them going beyond those encountered in the previous investigations. Some of these were concerned with the test materials and others were related to the development of a suitable procedure for handling the subjects. As regards the nature of the investigation, we may recall the fact, first of all, that we are dealing with a pedagogical phenomenon: the interpsychical contact between teacher and pupil. This relationship must not be disturbed by the test. The use of special test material and the appearance of strange test leaders should be restricted as much as possible, while the material as well as the explanation should be completely adapted to the interest and the perceptive abilities of the pupils. In short, the atmosphere of the class must not be disturbed and the teacher should remain the central element in the whole situation in spite of being so placed as to be invisible to the pupils. Although it was possible to adopt the test situation used in the former experiments, the age of the subjects in the new investigation made it necessary to introduce a different variety of test material better suited to the interests and stages of advancement of first- and second-grade pupils.

In the higher grades we used three different kinds of materials as targets. The subjects were allowed to choose on each trial one among five objects of the kind designated as the target material for that run (words, colors, or symbols), and they wrote their responses on a record sheet. In the present study we were directly confronted with the difficulty that the pupils of the first grade have not made sufficient progress in writing and that those of the second grade have not yet got command of the writing technique well enough to be able to write down their responses as an automatic reaction in such a way that the object of attention, the response to the target, remains primary. So we had to contrive test material in which no writing occurred. Furthermore, in order to facilitate the choice for these young pupils, we reduced the range of targets for each 12-trial run from five to three.

Of the kinds of targets used in the higher grades, only the colors were considered suitable for use in this study. Neither the arithmetical symbols nor the words used in the higher grades are sufficiently familiar to younger pupils to be suitable for this purpose. In order to keep the tests as much as possible in conformity with those of the higher grades, however, we did maintain the three-series session. For the first run we chose three geometrical figures,

triangle, square, and circle;

for the second run the objects,

tree, house, and boat;

while for the third run we used the colors,

red, yellow, and blue.

The procedure in the tests was almost identical with that of tests with higher grades (4) and it will therefore not be repeated in detail here. Briefly, it was as follows: The children were told by the person in charge of the test (the "test leader") that they were going to play a guessing game, and they were given a demonstration of how it was to be carried out. The teacher, who acted as agent (sender), sat at the rear of the room in a cubicle, screened from the pupils' view. Face-down in front of her were three cards bearing the targets for that particular 12-trial run.

In order to avoid having the pupils attempt to write, each subject was provided with a specially-designed record sheet. Instead of the usual column of spaces for noting down each guess, this record

sheet had 12 numbered lines on each of which appeared the three targets from which the subject was to make his choice. The targets appeared in different orders on different lines. The subject was to make his guess by crossing out one target on each line. At the top of the page were three unnumbered lines which were filled out for practice during the demonstration and which did not count in the final results.

When everyone was ready, the teacher referred to a list of random numbers to find which card was to be the target for that trial. The test leader did not know the order of the three cards lying in front of the agent. At a signal from the test leader, the teacher picked up the card, looked at it, and put it back down again. Meantime the subjects made their guesses for the first target. The test continued in this manner through the 12 trials of the run. When the trials for the first series (run) had been made, the sender's targets were changed and the second series was carried out; then the third. The three series were always done in the same order: geometrical figures, objects, and colors.

During the Amsterdam investigation the record sheet was given to the subjects folded in three lengthwise and stuck into an opaque paper cover in which a window the size of a line was punched out. The pupil saw only the line with three symbols that lay behind the window. For each trial the record sheet was pushed up one line higher in the cover, so that the next line became visible with the symbols in a different order.

The investigation made in the city of Dordrecht was slightly altered in one respect. Here attention was fixed especially on simplification of the actions the subjects had to make while filling out the answers. The question arose whether the concentration of the subject's attention on the line of targets, as in the Amsterdam work, might be considered satisfactory. It could be argued that rather than concentrating his attention on a few targets (three in the Amsterdam work) the subject could answer more intuitively if his attention were divided among a greater number. It is possible also that the larger number would be less likely to lead to patterning of calls. For this reason, the cover with the small window used in Amsterdam was discarded. Instead the record sheet was covered in such a way that a window the size of a whole series of 12 trials was open.

Results

Amsterdam Series

The Amsterdam investigation, in which I was assisted by Mr. Smolenaars, a graduate in psychology, included 26 classes in different schools in different quarters of the town. No special system of selection was applied. There were 12 first grades and 14 second grades, containing, altogether, 890 pupils, who made 32,040 guesses. The over-all results were insignificant (a deviation of +53, CR= 0.62). These general results are broken down in Table 1 to show the differences in scoring between boys and girls and between first and second grades.

Table 1

Results of First and Second Grades in Amsterdam

	No. of Calls	Hits	Dev.	CR
First-Grade Boys	7488	2447	− 49	−1.20
First-Grade Girls	6516	2274	+102	2.68
Total	14004	4721	+ 53	1.00
Second-Grade Boys	8784	2888	− 40	−0.91
Second-Grade Girls	9252	3124	+ 40	0.89
Total	18036	6012	0	
Total Boys	16272	5335	− 89	−1.48
Total Girls	15768	5398	+142	2.40
Grand Total	32040	10733	+ 53	0.62

Among the 890 pupils there were 438 girls and 452 boys. The girls made 15,768 guesses, of which 5398 were hits. Mean chance expectation is 5256. This gives a deviation of +142, with a CR of 2.40. The boys, however, scored 5335 hits in 16,272 guesses where M.C.E. would give 5424. This is a negative deviation of −89, with a CR of −1.48.

Although none of the CR's shown in Table 1 would be significant if corrected for group calling,[2] the best performance was supplied by the girls of the first grade, a phenomenon that is worthy of further investigation.

[2] See the footnote to Table 3.

As in previous tests (6), the last series of the test session have strikingly fewer hits than the first two series. (See Table 2.)

Table 2

DECLINE EFFECT WITHIN THE SESSION FOR THE AMSTERDAM TESTS

Series	No. of Calls	Hits	Dev.	CR
I	10,680	3637	+77	+1.57
II	10,680	3614	+54	+1.10
III	10,680	3482	−78	−1.61

Dordrecht Series

The Dordrecht investigation, in which I was assisted by Mr. Drogendijk, a teacher of psychology at a training college for teachers at Dordrecht, took place in 14 different first- and second-grade classes, with 544 pupils, 305 of whom were boys and 239 girls. A total of 19,584 guesses were made, with 7028 hits. According to M.C.E., 6528 hits was the most likely score. This gives a deviation of +500, a CR of 7.57 (Greville CR=6.90, P<.000001). Here too the girls scored higher than the boys. The CR for the girls was 7.09; for the boys, 3.84. Here, as in the Amsterdam tests, when the results were separated on the basis of grade, the higher scoring was done by the first-graders, CR=7.16, a highly significant result (corrected CR=6.44, P<.000001). The results of the second-grade subjects gave a CR of 2.64 (corrected CR=2.38, P=.02). Table 3 gives a survey of the results, grouped in different ways.

With these tests no striking difference between the results of the successive 12-trial series was found.

Combined Results, Amsterdam and Dordrecht

As the general design of the two investigations was the same, we presume that it is permissible to combine the results. In total, 51,624 guesses were made by 1434 pupils from 40 different classes with 40 different agents and 3 different test leaders. There were 17,761 hits, with a deviation of +553 and a CR of 5.16 (Corrected CR=4.64, P=.000003). The combined results are shown in Table 4.

Table 3

RESULTS OF FIRST AND SECOND GRADES IN DORDRECHT

	No. of Calls	Hits	Dev.	CR
First-Grade Boys	8064	2826	+138	3.25
First-Grade Girls	7416	2754	+282	6.94
Total	15480	5580	+420	7.16
Second-Grade Boys	2916	1024	+ 52	2.04
Second-Grade Girls	1188	424	+ 28	1.72
Total	4104	1448	+ 80	2.64
Total Boys	10980	3850	+190	3.84
Total Girls	8604	3178	+310	7.09
Grand Total	19584	7028	+500	7.57*

* When the significance of the Dordrecht work was tested by the Greville method, which takes account of the effect of multiple responses to the same target, this CR of the total was found to be 6.90. This represents a 9% reduction from the binomial CR, which agrees closely with the reductions that have been found in previous applications of the Greville method of analysis. It seems clear that the use of younger children as subjects and the reduction of the range of choice from five to three on each trial did not give any increased tendency to respond alike on a given trial. Thus it seems reasonable to accept the binomial CR's given throughout the paper as close approximations of the true values and to be cautious in stating that results are significant only in marginal cases in which a 10% reduction in the CR would fail to reach the usual criterion of significance. Similarly, in instances where the significance between two sets of data have been tested by a chi-square analysis applied to a 2 x 2 table, the value of χ has been reduced by 10% as the reasonable allowance for the "stacking effect" based upon many previous applications of the Greville method.

Table 4

COMBINED RESULTS FOR AMSTERDAM AND DORDRECHT

	No. of Subjects	No. of Calls	Hits	Dev.	CR
First-Grade Boys	432	15552	5273	+ 89	1.51
First-Grade Girls	387	13932	5028	+384	6.90
Total	819	29484	10301	+473	5.84
Second-Grade Boys	325	11700	3912	+ 12	....
Second-Grade Girls	290	10440	3548	+ 68	1.40
Total	615	22140	7460	+ 80	1.13
Total Boys	757	27252	9185	+101	1.29
Total Girls	677	24372	8576	+452	6.14
Grand Total	1434	51624	17761	+553	5.16

The difference between the total scores for Amsterdam and Dordrecht is significant ($P<.000001$), as the chi-square analysis in Table 5 shows. This finding points in the direction that simplification of the actions of the subjects and spreading of the attention over

the target in the new kind of record sheet might influence the results. A repetition of the tests with this new procedure might supply further information on the reason for the difference in the two cities.

Table 5

CHI-SQUARE TEST OF DIFFERENCE BETWEEN RESULTS OF THE TWO CITIES

	Amsterdam	Dordrecht	Total
Hits	10733	7028	17761
Misses	21307	12556	33863
Total	32040	19584	51624
% Hits	33.5	36.0	34.4

χ^2 (1 d.f.) = 30.70
χ (corrected) = 4.99
P < .000001

A chi-square test of the difference between the combined total scores for the boys and for the girls (corrected to allow for the group-calling effect, as pointed out in the footnote to Table 3) shows that the girls' scoring is significantly higher than the boys' ($P<.002$). This evaluation, however, was based on group performances and did not take into account the scoring tendencies of individual subjects. The combined results also show that the first grade gave higher results than the second. The CR for the first grade is 5.34, that for the second grade is 1.13. The corrected CR of the difference is 2.68, $P=.007$.

The clearly significant results of these investigations with unselected subjects confirm the results obtained with the higher grades and give, besides, an indication of a negative correlation between age and results.

DISCUSSION

An examination of all the tests I have made since 1951 in city primary schools with the teacher as agent shows that, so far, 4221

pupils have been tested in 127 different classes with 127 different teachers as agents and ten different test leaders.

The fact that the results of such a great number of subjects in different schools in both Holland and America yield significant results with striking consistency shows ever more clearly that an extrasensory factor is active in the teacher-pupil relationship. What factor that is has not yet been investigated in these tests. Some possibilities of what may be occurring are these: The pupils may be borrowing their knowledge by means of telepathy from the teacher's consciousness. The teacher may be forcing her conscious thoughts upon her pupils. Or the pupils may be getting the target directly by clairvoyance. Since it is difficult to separate the different aspects of extrasensory perception, psychokinesis, and precognition, the Greek letter "psi" is used to cover them all as a collective name. In this special case we cannot make out absolutely whether any sort of distinctions can be made.

That the problem is very complicated is certainly proved by the previous experimental reports in this series of tests in which the agent was a stranger to the pupils and to a high degree by the tests of Anderson and White (2) and Anderson (1), which showed a positive correlation between the results of clairvoyance tests and the attitude of teacher and pupil toward each other at the primary and secondary school level in the U.S.A.

The fact that my studies as a whole have yielded significant results from elementary pupils when their own teachers acted as agents should not cause us to lose sight of the striking variations in performance which did occur. Investigation of the reasons for differences in scores under different conditions may help to throw further light upon the nature of the extrasensory factor in question. The present investigation yielded three such differences which deserve to be considered as fully as the presently available facts will allow and which should be kept in mind for further study.

One is the difference between the boys and girls. In the earlier work with fifth- and sixth-grade pupils there has not been a striking difference between the sexes, but the boys alone have been significant in two investigations (one in Holland and the other in America) while the girls were significant in the remaining study, done in Holland. None of the sex differences at the higher grade level

were statistically significant, but the significant results of the chi-square test of the difference between the boys and the girls in the present study, it seems safe to say, reflect some genuine factor or factors favoring the girls over the boys even though I cannot suggest what the cause might be. It will be interesting to see if this sex difference in the school room is found in later studies at the first- and second-grade levels.

A second noteworthy difference produced by the present experiment is that between the first and second grades. In this instance the first grades of the two cities produced a higher level of performance than the second grades, and the difference is of a magnitude that is difficult to attribute to random variation. It is, again, only possible at this stage to speculate upon the reason for this difference. Might it indicate that the experimenters, in the effort to make the level of the instructions and procedure simple enough for the first-grade children, actually made the test too elementary to appeal to second graders? At least the difference, along with the others noted, shows how subtle and elusive the extrasensory factor in the schoolroom really is.

The third and most striking difference is the one already pointed out as having occurred between the two cities. To what might it be attributed? I have already mentioned the possibility that the difference between the test procedures, the use of the small window for each trial in the Amsterdam work, as contrasted with the large opening which exposed all 12 trials at once at Dordrecht, might have been a contributing factor. There may also be other factors that influenced the results, such as the unavoidable personal difference between the manner in which the tests were presented by the different test leaders in the two cities, the difference between large and small cities, and the difference between classes. Without further analysis we cannot draw any final conclusion.

References

1. Anderson, Margaret. Clairvoyance and teacher-pupil attitude in fifth and sixth grades. *J. Parapsychol.*, 1957, **21**, 1-12.
2. Anderson, Margaret, and White, Rhea. Teacher-pupil attitudes and clairvoyance test results. *J. Parapsychol.*, 1956, **20**, 141-57.
3. ———. A further investigation of teacher-pupil attitudes. *J. Parapsychol.*, 1957, **21**, 81-97.

4. VAN BUSSCHBACH, J. G. An investigation of extrasensory perception in school children. *J. Parapsychol.*, 1953, **17**, 210-14.

5. ———. A further report on an investigation of ESP in school children. *J. Parapsychol.*, 1955, **19**, 73-81.

6. ———. An investigation of ESP between teacher and pupils in American schools. *J. Parapsychol.*, 1956, **20**, 71-80.

Kleverlaan 151
Haarlem, The Netherlands

EXPLORING THE PARAMETERS OF RESEARCH VARIABLES[1]

By GERTRUDE R. SCHMEIDLER

I will propound a heresy.

One of our pious generalizations has long been that parapsychology needs a repeatable experiment—and I suppose this means that my proposal is an impious generalization. It is that what laboratory parapsychology needs is long series of *non*-repeated experiments.

The key word here is "series." According to this prescription for our mild case of retarded development, any follow-up project ought to explore some new aspect of what has preceded it; it should attack the original problem from a new angle. There should be a method in our work; but it should not be the method of tedious attempts to duplicate what has been done already.

For example, let us suppose that we are ready to begin experimenting because we have an idea about a method that will give high ESP scores. We plan our procedure, perhaps run a few practice subjects to make sure that it goes smoothly, and then test a predetermined number of subjects for the formal experiment. Our idea works out; ESP scores are high. Good. We write this up for publication; but what do we do next? My recommendation is that we should *not* try to do exactly the same thing again to make sure of it (though it might be wise to have someone else try to do exactly the same thing, to see if it was our enthusiasm or our personal charm that resulted in the first high scores). Instead we should modify what we believe was the key part of the procedure so that we expect the ESP scores to be lower. Or modify it so that we expect them to be even higher—it doesn't matter which. If this too comes out as planned, we should change the procedure again in a way we think will bring forth intermediate scores. Or still higher ones. Or still lower ones (though of course we would not direct all the changes toward lower and lower scores). I submit that if each of four or five

[1] Presidential Address presented on Sept. 11, 1959, at the second annual convention of the Parapsychological Association in New York City.

such systematic variations shows the expected changes, this will give us far more useful information than four or five successful repetitions of the original technique.

Now, you may say, this would be all very well except for one big defect. Critics of ESP demand that we show them a repeatable experiment; and until we demonstrate repeatability we have not met their attack. That is true. But even if our primary object in doing research is to help public relations (and I suspect that few experimenters have this as their primary object) we can probably achieve it faster by advancing our own knowledge than merely by taking the defensive in a continuing, shifting argument. Die-hard critics of parapsychology are hydra-minded; as soon as one of their points is answered they make two new ones. I refer you here to the "fact" that the physicist Max Planck tells us he has learned. "A new scientific truth," he says, "does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it" (12).

And consider the advantages of exploring the range of an independent variable (which is a dignified description of my recommended method) instead of testing and retesting it in a narrow part of its range. The first, to be mentioned only in passing, is that the experimenter is less likely to grow bored with what he is doing—though to one who has been bored herself, and felt that her subjects' interest lessened correspondingly, this seems no small item. Of the other benefits, I shall state quickly two if the experiments work out as we expect them to, and then spend a little more time on others if the experiments do not work out.

First, if the results come out as planned. It has often been said that the best test of a theory—and the most severe one—is its ability to predict new facts. When our first bright idea has succeeded, we have met this test once. It is obvious that if we replicate the method, we make no new predictions. However, if we change the procedure, we do.

Next, if we can show that ESP or PK scores vary systematically with changes in some other condition, we have taken a long step toward following Dr. McConnell's sage advice of last year, to scale and quantify our observations so that we can measure effectively

what we are working with (9). As Lord Kelvin put it more informally, then we "know something about it."

But the advantages will be more spectacular if our predictions fail. Let me elaborate on this with an example.

Suppose that no one had ever correlated ESP scores with intelligence test scores. And suppose that you were wondering about the basic nature of ESP ability, and began to think that the two essential things about it were (a) a sensitivity to subtle, faint cues that most people seldom notice, and (b) a readiness to respond to those cues even when they are incongruous with expectations or habits. If you have gone this far, your next step might be a sudden, shiny insight. You might say: But this is exactly what bright people do when they solve problems successfully. Intelligent behavior is also essentially a matter of observing inconspicuous or incongruous stimuli and of putting them to use. Maybe intelligence and ESP are variants of the same process! Let's put this to the test, you might say, by correlating them.

You go ahead and try it. Using the college students that are at hand, you give them ESP tests, at the college. You then administer intelligence tests, or with the permission of the Admissions Office you take your subjects' old intelligence test scores from the files. You correlate. Let's suppose you find a positive relation as Humphrey did (7), or as the Vasses did with first-grade children tested by their own teacher (19). What next, after you write up your findings? There would be no harm in doing it again; but there would be little interest, with similar subjects in a similar setting, in getting similar data. Instead you attack the same theory in a different way. You select, the following year, the brightest students in your classes. You expect high ESP scores from them. You might even tell them frankly, as a recruiting device, that they are selected because their known high intelligence should aid them in doing well at ESP. You test them; and sure enough, their average score is higher than the average of last year's unselected subjects.

What next? Again, there would be no harm in doing it over. But if you hold to this heresy of mine about the non-repeated experiment, your next step ought to be to find still another approach to the same theory; to explore the independent variable; to change conditions so as to see if you can make the average ESP score drop

as you have already made it rise. Thus as a well conducted heretic you might go to the other end of the intelligence scale. Perhaps, like Esther Bond Foster many years ago, you would run your next series at a school for retarded children. Your theory predicts that their ESP scores will be low. The teacher administers ESP tests to them; or if you happen to like retarded children and get along well with them you administer the tests yourself. Then, if your results are like those of Bond (3) or like those of Neureiter that Ehrenwald defends and explains (6), you will find their ESP scores are well above chance expectation.

Now this is the real benefit of changing procedures. You have learned something that it could have taken a long while to find if you stayed with the original successful method. Your first good results stand—they're facts, and there's no denying them—but your first theory of the near-identity of intelligence and ESP ability has crumbled. You have outgrown your original approach; it is time to salvage some pieces of it and construct a new thory that will reconcile all your facts. You might begin to speculate, for example, that though the utilization of incongruous cues is indeed typical of both high intelligence and ESP, it is not exclusive to them; that it is also a characteristic of perceptual flexibility as opposed to rigidity; that a factor common to both the bright students at college and the retarded children in a happy classroom was that they felt at home in their environment and therefore, probably, permitted themselves freedom (or flexibility) within that environment; and thus that the variable related to ESP success may have been flexibility in the use of unexpected cues, the result of feeling accepted and at ease.

It is tempting to pursue this line of thought, because at present it is one of my own pet theories. In fact, one of my students is now testing children in an institution and in elementary school for the rigidity or flexibility of their perception, and is also administering ESP tests to them. But it has little to do with our present topic, and I will put it aside for now (but return to it later).

The next on the list of advantages of non-repeated experiments should particularly attract the economical research worker. If our initial experiment gives results different from what we expected, we need not (using this approach) consider that it was a waste of effort. Instead it can represent the base line with which later results in the

series are to be compared. After we have brooded over our unexpected findings, and perhaps talked about them with our nearer friends or written about them to our more distant ones, we may well think that we know where the trouble lay. If then we change the procedure in a second experiment, and find the expected change in the data, both the earlier scores and the later ones have become meaningful.

At last year's Convention we heard a neat report of just such a modification of a failure that turned into a success. Van Busschbach had expected first- and second-grade children to do well at ESP, but he found on testing them that their ESP scores were low. Instead of dropping them from his later work, he ingeniously changed their method of response so that it was more concrete and easier. The improvement in ESP scores was striking (**18**). Surely this difference in scores is more interesting than a mere statement of high scores would have been. And I think we should all be grateful to him or to any other experimenter for some systematic exploration of the effect on ESP hits of varying the concreteness and difficulty of tasks at any age level. (We would expect the easier, more concrete tasks to give high ESP scores for young subjects and for stupid subjects, wouldn't we?——but probably to lose interest quickly for mature, intelligent subjects, and thus show a rapid decline effect; whereas more complicated and varied tasks might be too much for young or stupid subjects to cope with, but be challenging for brighter subjects, and thus might show a positive relation to intellectual level. But this is off the track; I keep being drawn into the content of parapsychology when my paper is supposed to be about method.) In any event, the way to study concreteness and simplicity of task, if that is our object, would be to run not only one or two tests, provocative as these may be, but instead to run a longer series which covers a wider range.

My argument so far is that in a first experiment neither low scores nor high ones should satisfy us and that either should be regarded as a proper starting point for further research. However, an equally good starting point could be scores which are intermediate or which seem contradictory. Here, beginning in the middle, we would hope to push out to both ends. Cadoret (**5**), for example, comparing subjects who had and had not taken drugs, reports that

after dexadrine his subjects had lower ESP scores at card-guessing but higher ESP scores at picture-drawing. He also, in summarizing earlier work, reminds us that after a drink a Dutch subject had higher ESP scores than before, but that a subject at the Duke Laboratory had lower ones. These odd items might persuade the timid that dosages of dexadrine and alcohol were, above all others, the conditions *not* to investigate. But Cadoret's theory and Reeves' somewhat different theory (**14**) of such effects indicate that, on the contrary, research might quickly tidy up the whole problem. If, as these two independently suggested, the drugs affect motivation and the motivation affects ESP scores, then if drugs are varied *and motives are measured,* the intermediate, contradictory data provide an excellent beginning for systematic investigation of two interacting variables.

Perhaps another bit of exploratory research should be mentioned in this context, even though it is not concerned with ESP. It deals with the TAT, a test of personality in which the subject is shown ambiguous pictures and asked to tell imaginative stories about them. It was expected that a short time after drinking whiskey the subjects would tell longer, livelier, more rambling and more revealing stories than usual. The first experiment on this topic was performed by an investigator who wanted to be sure that no breath of scandal would attach to him for making his subjects drunk. He succeeded both in making them drunk and in keeping university approval by following an orthodox procedure. His subjects came to a physiological laboratory; the experimenter, in a starched white coat, measured a stiff drink into a laboratory beaker; he had the subject drink it; he waited quietly for an interval so that the alcohol would take effect; and then, following the clock and the rule book, he asked the subject to tell imaginative stories. The stories that followed were short and prosaic; and this seems natural, because the subjects were half-asleep. In fact, the results are easy to understand. We all know that alcohol is a depressant.

However, a second experimenter was not satisfied with this post hoc explanation. In a technique which I emphatically do not recommend to you (and which so far as I know he has never published) he made appointments for the late afternoon, greeted his subject in an office furnished like a living-room, waved him into

a comfortable chair, and then, varying his comments according to the weather, said that it was too nice a day—or too disagreeable a day—to work; they should have a drink instead. Glasses had previously been obtained which held the appropriate amount. Alcohol was measured simply but lavishly by pouring as much as a glass would hold. Subject and experimenter had their drink and chatted convivially—sometimes uproariously, according to friends in nearby offices. Eventually the experimenter told his drinking companion that they had better get some stories told, reached for the TAT cards, and administered the test. It probably will not surprise you that the stories told to the ambiguous TAT cards under these conditions were as lively and uninhibited as the stories in the laboratory had been prosaic and dull, even though the alcohol intake and the delay after drinking it were nearly the same.

How much more complicated this makes the research problem! It looks as if the subjects in the first experiment were not responding to alcohol; they were responding to alcohol-administered-in-a-laboratory; and subjects in the second experiment were not responding to alcohol, but to alcohol-taken-informally-with-a-friendly-man. Cadoret's subjects might have been affected, not by dexadrine, but by dexadrine-in-a-laboratory-followed-by-picture-drawing and by dexadrine-in-a-laboratory-followed-by-card-guessing.

Of course we all know about this dependence on the specific situation already, for in one of the classic experiments of parapsychology Pratt and Price showed that children were responding, not to ESP cards, but to ESP-cards-administered-by-a-particular-experimenter (13). It is similar too to one of the often repeated experiments of modern psychology, which showed that of the politically conservative students who read a passage recommending changes in capitalistic practices, those who knew the author was the head of the National Association of Manufacturers usually approved of it, but those who were told falsely that it was written by a socialist usually condemned it (2). We can be sure that very often both the subject's background of motivation and his estimate of the total situation interact with the official stimulus (as the experimenter has defined it). It is not the official stimulus alone which produces the results.

Let me put the same statement in a more outrageous form. I have been saying, quietly, that the experimenter often does not know what his own procedure is, or what the stimuli are, because he does not know how they look to his subject's eyes. If so, it is only the greatest good luck if the experimenter reports his own research correctly, or if he knows what interpretation to make of his data. How important it is, then, for him to keep varying his research approach, using different kinds of subjects and different procedural techniques, when it is so easy for him to be misled about his independent variable.

These accusations of mine may seem radical ones which strike at almost all of psychology and parapsychology, but they are not novel. They represent, as you know, the basic thesis of Murphy's field theory (**10**), which emphasizes the interaction of social and biological factors with each subject's past history and motivation, to show that no one of them is ever the single determinant of a psychological event. Thus another and probably the greatest advantage of systematically varying the procedure in a series of experiments is that in this way we have more opportunity to explore the field conditions of our research and to get an inkling of what special factors or background conditions, which we had previously taken for granted, were crucial to our findings.

Perhaps another way of thinking about the same problem is as a difficulty in semantics. The words we use are seldom sufficiently precise to specify either our procedure or our ideas. But if we perform different operations, all of which can be described by the same words, then if we get the same results in all cases we will know that those broad, general terms were appropriate, and if we get different results we will know we must put in special modifiers. Let me return, to give an example of this, to the flexibility issue which I raised earlier.

Suppose that my student finds that those children whose perception is more flexible tend to have higher ESP scores than those whose perception is more rigid. The next research might use the same procedure on other subjects expected to be still more rigid or more flexible; and this is the sort of quantitative variation of the independent variable that I have been advocating. But a more urgent question for our theory is whether the general term "flexibility,"

with all its connotations and applications, is the proper one to apply to these limited findings. I would therefore advocate even more strongly a follow-up project using different forms of perception, or learning and thinking, or social behavior, to find if there are qualitative variations in the meaning of "flexibility." For example, will drugs which narrow the field of attention or of vision also reduce ESP scores? They should, according to the first broad statement. And what of drugs that make the subjects' responses more labile and variable? They should increase ESP scores, according to this simple theory; if they do not, the theory must be modified.

Or take the case of learning. We are quite sure, now, that either chronic anxiety or temporary anxiety caused by failure results in faster learning by conditioning, but in slower learning of the trial and error (or thoughtful) type. This is often interpreted to mean that anxiety lessens flexibility but at the same time increases the subject's tight concentration on specific response patterns—the same sort of specific response patterns that Rhine has associated with psi-missing (**15**). We can put these two interpretations together and predict that when a subject's level of anxiety is so high that he conditions quickly but solves problems badly, he should show more psi-missing; when a subject feels so little anxiety that his conditioning is unstable but his problem-solving is good, his ESP scores should be higher.

Or take social attitudes. Prejudice, when associated with high scores on the F-scale or other measures of authoritarian tendencies, is often interpreted as a need for dependence on rigid authority, and thus for rigid prejudgments. These help a person feel secure but also make him give inappropriate responses to people or ideas. Subjects with marked racial prejudice and with high F-scale scores are thus presumed to be rigid and inflexible. If they are rigid, they should by this theory show psi-missing. But would they? Or would their ESP scores be high if the ESP test were administered by an authority figure who made stern demands on them for ESP success?

Only by such wide-ranging experiments, which cover diverse fields to which the same labels (like rigidity and flexibility) apply, can we judge whether those labels are appropriate. No broad theory can be tested adequately by replications of a specific method.

In summary, I am recommending that instead of repeating our experiments we should use three methods of serial experimentation: (1) If we are investigating a precisely defined variable (such as IQ or chronological age) we should examine a wide range of that variable. (2) If we are dealing with a concept which is not precisely defined (such as flexibility, spontaneity, or selflessness) we should test a wide variety of situations to which that concept has been applied. (3) We should routinely use different kinds of experimenters, different techniques of ESP testing, different populations of subjects, and so on, before we generalize with any confidence about data from the arbitrary procedures with which we began.

(None of this, of course, should be taken to mean that any experiment we perform must be short and simple, like the annals of the poor. Each unit of the series may be a long-term, well-designed project which itself examines many conditions and tests many subjects. I am not asking you to forget the importance of levels of significance.)

Having at last stated my argument, I will call in some support for it from two articles which came my way in the last few weeks. The first is by Neyman and Scott, who were awarded the Newcomb Cleveland Prize at the A.A.A.S. Convention in 1958 and who presented there a stimulating paper on Stochastic Models of Population Dynamics. They develop a theoretical model for research in diverse fields, and make the delightful statement in its defense that it can serve two purposes, "one esthetic and the other utilitarian." It is their esthetic purpose, not their stochastic model, which I will quote to you. They continue: "When it happens that a number of apparently disconnected manifestations of a class of phenomena are logical consequences of a single simple combination of a few 'elementary' mechanisms then, to most of us, this circumstance is a cause of distinct esthetic satisfaction" (11). And it is a pleasure, just as they say, isn't it, to find that separated facts fit together? It is a pleasure that repeated experiments cannot yield, because with them we either find the same fact over and over, or else fail to find it. You make it possible to obtain this pleasure, or to achieve this esthetic purpose, only by making sure that the diverse facts are available. And you do this by using diverse procedures.

Bruner, a psychologist, makes a similar point without using the same terms. He says, "If it is a good theory . . . it should permit us to go beyond the present data both retrospectively and prospectively. We go backward—turn around on our own schemata—and order data that before seemed unrelated to each other. Old loose ends now become part of a new pattern. We go forward in the sense of having new hypotheses and predictions about other things that should be but have not been tested. When we have finished the reorganizing by means of the new theoretical coding system, everything then seems obvious, if the thing fits" (**4**). Note those "loose ends" that are needed for the new and better theory. Repetitive research is not likely to provide them; but if every test of a theory is also an exploration into what Bruner calls "other things that should be but have not been tested," there may well be enough loose ends for all of us to do our theorizing with. And if our theory was sound, we will have the esthetic satisfaction of seeing that "the thing fits."

You may have been wondering if I would neglect the utilitarian purpose that Neyman and Scott mention. Their point is that a theoretical model which works can be put to use. This fits in with what Bruner too has been saying: that the thing to do with a good theoretical model is to use it for new predictions. If its predictions are effective, it serves its utilitarian purpose. Thus the recommendations for other fields of science support this thesis that a series of predictions should be made and tested to examine the validity of a theory.

In a way, this is only what other experimenters in parapsychology have been doing all along. When one of Soal's brilliant subjects obtained outstanding extrachance results with one agent, Soal varied the agent (**17**). When Anderson and White found repeated significant differences in ESP score relating to the attitude of student and teacher toward each other, in grade school and high school, they began testing in reform school and college (**1**). When Scherer's subjects, restricted to a maximum of two or three guesses a day, built up high averages on his marble machine, he used the same machine with the requirement of many responses (**16**). All these investigators found differences in results between the first

method that worked and later ones that did not work so well; and those differences are essential to an understanding of their work.

I will conclude by saying what you have probably all noticed. In my argument that experiments should not be repeated, there is really no contradiction of the older thesis that they should be repeat*able*. You can therefore properly accuse me of overstating what I have to say. In defense, let me quote Stephen Leacock, who ends an article on a very different topic with these remarks, "All this, you will say, is exaggerated, is over-colored, is not truth. Very likely. But a half truth in argument, like a half brick, carries better. High colors show up where neutral tints blend to nothing, yet the main truth gets over" (**8**).

You are all familiar with the method of seeing faint stars at night: if you look a little away from them, you can see them best. Perhaps in a somewhat similar way we will be most likely to find the repeatable data and the effective predictions if we do not direct our procedure at monotonous repetitions, but instead aim first at one side and then at the other of our initial findings, exploring quantitatively and qualitatively the variables we are studying, and at the same time subjecting our theory to a series of tests of its predictive powers.

References

1. Anderson, M., and White, R. A survey of work on ESP and teacher-pupil attitudes. *J. Parapsychol.*, 1958, **22**, 246-68.
2. Asch, S. E. *Social Psychology*. New York: Prentice-Hall, 1952.
3. Bond, E. M. General extrasensory perception with a group of fourth- and fifth-grade retarded children. *J. Parapsychol.*, 1937, **1**, 114-22.
4. Bruner, J. S. Going beyond the information given. *Contemporary Approaches to Cognition*. (A Symposium held at the University of Colorado). Cambridge: Harvard Univ. Press, 1957.
5. Cadoret, R. J. The effect of amytal and dexadrine on ESP performance. *J. Parapsychol.*, 1953, **17**, 259-74.
6. Ehrenwald, J. Telepathy and the child-parent relationship. *J. Amer. Soc. psych. Res.*, 1954, **48**, 43-55.
7. Humphrey, B. M. ESP and intelligence. *J. Parapsychol.*, 1945, **9**, 7-16.
8. Leacock, S. Education gets longer and longer. *The College Years*. (A. C. Spectorsky, Ed.). New York: Hawthorne, 1958.

9. McConnell, R. A. Scaled measurement in psi research. *J. Parapsychol.*, 1958, **22**, 237-45.
10. Murphy, G. *Personality: a Biosocial Approach to Origins and Structure.* New York: Harper, 1947.
11. Neyman, J., and Scott, E. L. Stochastic models of population dynamics. *Science*, 1959, **130**, 303-308.
12. Planck, M. *Scientific Autobiography and Other Papers.* New York: Philosophical Library, 1949.
13. Pratt, J. G., and Price, M. M. The experimenter-subject relationship in tests for ESP. *J. Parapsychol.*, 1938, **2**, 84-94.
14. Reeves, M. P. A topological approach to parapsychology. *J. Amer. Soc. psych. Res.*, 1944, **38**, 72-82.
15. Rhine, J. B. The problem of psi-missing. *J. Parapsychol.*, 1952, **16**, 90-129.
16. Scherer, W. B. Spontaneity as a factor in ESP. *J. Parapsychol.*, 1948, **12**, 126-47.
17. Soal, S. G., and Bateman, F. *Modern Experiments in Telepathy.* New Haven: Yale Univ. Press, 1954.
18. van Busschbach, J. G. An investigation of ESP in the first and second grades of Dutch schools. *J. Parapsychol.*, 1959, **23**, 227-37.
19. Vasse, C., and Vasse, P. ESP tests with French first grade school children. *J. Parapsychol.*, 1958, **22**, 187-203.

College of the City of New York
New York, N. Y.

SOME PROBLEMS IN GROUP TESTING OF ESP

C. J. Adcock and D. Quartermain

ABSTRACT: In these exploratory tests four groups of college students were tested by three different ESP card techniques (known as the BT, DT, and GESP methods). The total scores in these twelve different sections of the data gave a pattern of both negative and positive deviations that varied too much to be reasonably attributed to chance. Twelve "control" sections, however, showed no extrachance variation. The subjects were classified as to whether they believed in the possibility of getting significant scores (sheep) or not (goats). There was no indication that these attitudes affected the scoring. The authors suggest that the statistically significant negative scoring for the BT method may be due to skepticism combined with a decline effect.

The two authors of this paper are members of the Psychology Department at Victoria University of Wellington, New Zealand.—Ed.

Introduction

This experiment was designed with three objectives in view:

(1) to test for ESP in the ordinary college classroom setting;

(2) to compare the results with those obtained from a set of control targets not aimed at by the subject;

(3) to inquire into the reason for significant negative scores, if they were obtained, from those subjects who were skeptical of the possibility of ESP.

Procedure

Four groups of first-year psychology students (numbering 18, 28, 28, and 27 respectively) were given a short talk on ESP and were then given several group tests using conventional ESP cards in shuffled, "closed" decks. Three runs of 25 trials were made—the first run using the DT technique; the second, BT; the third, GESP. Two different experimenters acted as agents, or senders, in different parts of the GESP tests. The agent stood on a platform before the subjects, with a long screen four feet high between him and them. Behind this screen he shuffled the cards and held up each one with the back towards the subjects as they made their guesses. Some weeks after the administration of the test the written records of the subjects were scored and the results were checked (by Dr. G. L.

Mangan). The score sheets were then rescored, using in place of the original targets a new random arrangement of the target cards produced by shuffling the pack several times. The subjects were quite unaware of the intention to re-score their calls in this second series, so it is unlikely that the second target order would have served as a second set of ESP targets. In this way we hoped to have the equivalent of a control group where we would expect only chance results.

After the ESP tests were described, but before they were taken, the subjects were asked to state whether or not they believed in the possibility of significant results. On this basis they were divided into "sheep" and "goat" groups.

Results

Deviations from chance expectation within each group were evaluated separately for each test method, using the Greville formula which allows for the stacking effect due to the group guessing for the same targets. These data are presented in Table 1.

Table 1

Deviations from Mean Chance Expectation

Group		Expected	DT	BT	GESP	Total
A	ESP Targets	90	7	− 4	− 7	− 4
	Control Targets	90	− 8	− 2	11	1
	S.D.		8.67	8.16	8.77	
B	ESP Targets	140	8	−28*	− 9	−29
	Control Targets	140	6	−13	−10	−17
	S.D.		10.39	11.33	11.05	
C	ESP Targets	190	−29*	−18	24*	−23
	Control Targets	190	19	10	−11	18
	S.D.		13.69	12.13	12.22	
D	ESP Targets	135	10	−12	− 9	−11
	Control Targets	135	9	1	13	23
	S.D.		12.15	9.41	11.32	
Total	ESP Targets	555	− 4	−62**	− 1	−67
	Control Targets	555	26	− 4	3	25
	S.D.		22.62	20.76	21.83	37.67

* Significant at .05 level by the Greville method.
** Significant at .01 level by the Greville method.

Examining the scoring for each method, we find that, for the actual target, three of the separate section totals show a deviation which is significant at better than 5%. With 12 possibilities, we would expect one result at the 5% level as reasonably likely (actually one in 20) but three or more such results would occur by chance less than twice in a hundred cases.[1] It would appear that the over-all results are not due to chance fluctuations.

If we combine probabilities by the Fisher method (entering positive[2] probabilities as corresponding negatives; i.e., 5% positive =95% negative) we get an over-all probability of about .01 (χ^2= 26.02; P<.011). This confirms that we have here extreme values (negative and positive) significantly greater than could be expected on a chance basis.

It should be noted that this does not mean that the combined results show significant negative or positive scoring. We know that this is not so. What the chi-square test indicates is that the section deviations (the 12 group-condition scores) are too extreme to be explained on a random basis and that we have to look either for something which will explain reversal of direction of psi capacity or independent explanations of positive and negative scoring. If the former could be successfully achieved, we would be in a position to predict the direction of scoring. Under such conditions all our significant scores would accord with hypothesis and so be combinable as positive values.

By way of contrast, our control scoring on the reshuffled packs gives us not a single section total which is significant, and the grand total has a probability of .50. It does appear that something is operating in the one case which is not present in the other.

[1] P=.0197 by the formula $_nC_r\, p^r q^{(n-r)} = \frac{n!}{(n-r)!\,r!}\, p^r q^{(n-r)}$

Where n=number of trials
r=number of trials significant
p=significance level, here .05
q=1−p.

Since we require the probability of r or more trials being significant, calculations for each value of r must be made separately and the probabilities added. In practice, allowance needs to be made for only two values above r, as higher values make no appreciable difference. In the present example the value for r alone was .0173 and this would have been sufficient indication of significance for most purposes.

[2] Since negative deviations predominate, they must be taken as standard.

It would appear from our results that we have evidence of both positive and negative psi effects. To explain this we must either have separate hypotheses for the opposing results or find some qualifying factor which would determine whether psi should operate in its positive or negative form. We had anticipated this and had made provision to test the hypothesis that positive (sheep) or negative (goat) attitudes towards ESP might result in positive or negative scoring. Table 2 summarizes the relevant data.

Table 2

DEVIATIONS FROM MEAN CHANCE EXPECTATION IN TERMS OF SUBJECTS' ATTITUDE TOWARD THE POSSIBILITY OF ESP

Group	Attitude	Expected	DT	BT	GESP	Total
A	Sheep	40	+ 2	0	+ 5	+ 7
	Goats	25	+ 2	− 6	− 6	−10
B	Sheep	35	− 4	−10	− 1	−15
	Goats	95	13	−14	− 4	− 5
C	Sheep	105	−12	−18*	+11	−19
	Goats	85	−17*	0	+13	− 4
D	Sheep	60	− 3	− 6	− 2	−11
	Goats	65	13	− 1	− 2	+10
Total	Sheep	240	−17	−34*	+13	−38
	Goats	270	+11	−21	+ 1	− 9

Significance levels have been evaluated here without allowance for the stacking effect since they are not important for this study and therefore no rigorous test is necessary. The direction of the deviation itself is decisive.

* Significant at P = .05

Two section scores are probably significant at the .05 level, one for the "sheep" and one for the "goats." In both of these cases the results are negative. One method total for the sheep is also significantly negative at the same level, but this figure is not independent of the lowest section score for the sheep. Considering group totals we find that in one case (A) sheep are positive and goats negative, in another (D) the position is reversed, while the other two groups show negative results in both cases. The grand total shows negative deviations for both sheep and goats but a rather higher negative

total for the sheep. The difference, however, is not significant. Degree of belief is evidently not a key to the solution in this case.

Discussion

We now need to ask what other hypotheses could be advanced to explain our results. A study of the breakdown of the data suggests some clues. It will be noted from Table 1 that only one of the three methods gives us a total deviation which is significant. This is BT whose deviation is significant at .01. For this condition all group deviations are negative in direction, although only one is individually significant. By contrast, GESP shows an insignificant total deviation of −1, with one group deviation being significant at .05 in the positive direction. This might be explained as a differential effect of skepticism, telepathy appearing to these doubting students as rather more credible than clairvoyance.

This hypothesis alone would require that DT should also give negative results. As in GESP, the total DT deviation is again insignificant, but one section total is negative and significant at less than the .02 level, and this is as would be predicted. However, one might expect DT to provoke more skepticism than BT. That it has not done so may possibly be explained by the operation of a second factor, positional effect. Rhine (1) shows that there is a tendency for negative influences to develop as a session proceeds, positive results frequently becoming negative during the second half of a session. Perhaps the failure of the decline effect to show up in the third part of the session may be attributed to the smaller degree of skepticism for the GESP method. In other words, the tendency toward psi-missing due to the decline effect may have been cancelled out by the favorable effect of using the GESP condition. But these are, as indicated, only clues as to what might have caused the scoring trends in the present data.

The two factors we have considered here *could* explain the pattern of our results. Other possible factors discussed by Rhine seem less likely as explanations here. Personality variables such as extraversion-introversion, expansion-compression, etc., would be unlikely to have the differential effect manifested by our group totals since there was no systematic assignment of students to the four classes and we may therefore assume that personality factors were

evenly distributed over the four groups. Systematic checking errors would similarly be contra-indicated. Both types of factor might be operating but masked in effect through the cancelling of negative and positive deviations. This could be detected only by appropriate division of our subjects. This is not possible at this late stage but neither is it important in relation to our present data since we have a significant result and it is this which we have to explain. The trend which is manifest in our data seems best explained in terms of the hypotheses advanced. A more rigorous and comprehensive approach will be necessary to confirm and supplement these hypotheses.

Conclusions

1. We have shown that our results, although giving a pattern of both negative and positive deviations when considered by sections, cannot be reasonably explained as due to random effects.
2. Our procedure of scoring the subjects' calls against an arbitrary target did give a pattern of random deviation as predicted.
3. The general trend of deviations in the experimental data is in the negative direction.
4. The most consistently negative scoring occurred for the BT condition, and this was statistically significant. It is possibly explicable as a combined effect of skepticism (which could be expected to be higher for clairvoyance than for telepathy) and a decline effect.

Reference

1. Rhine, J. B. The problem of psi missing. *J. Parapsychol.*, 1952, **16**, 90-129.

Victoria University of Wellington
New Zealand

INFLUENCE OF THE POSITION OF THE SUBJECT IN RELATION TO THE TEST CARDS UPON ESP RESULTS

By Rémy Chauvin[1]

ABSTRACT: The subjects in this clairvoyance test (one adult and seven children) called targets laid out in a variety of positions: in front of the subject, behind him, overhead, below, or at an angle. The targets were the numbers 1 and 2 written on cards and enclosed in opaque envelopes. These were arranged in five rows of six columns each so that any evidence of ESP in the distribution of hits might be more easily spotted than in the more commonly used arrangements of the target cards. The total number of hits scored by the eight subjects did not depart significantly from chance. The analysis of the hit distribution, however, showed a negative score for the three left-hand columns in the card arrangement, and a positive score for the three right-hand columns, the difference between the two being significant. The author points out that even though there were marginally significant differences indicating that position of the cards in relation to the subject affected the results, these differences may have been due to psychological factors.

Prof. Chauvin is Director of the French National Bee Laboratory and the Laboratory of Ethology.—Ed.

Introduction

Problem

Until the present time, experimenters in ESP have not found any evidence from their general observations of tests carried out with ESP cards that the position of the subject relative to the cards influences the results. But the cards such as those used in parapsychological experiments have generally been arranged in a limited number of ways: the pack has been left intact for "down through" calling; or the cards have been removed one by one as they were called.

The possibility remains of arranging the cards in several rows for the subject to call in a regular manner. If ESP is influenced by the subject's position in relation to the cards, this fact might be easier to discover by such a method, which would reveal any unusual

[1] I wish to thank Prof. J. B. Rhine and Dr. J. G. Pratt for their helpful suggestions, especially in regard to the analysis and evaluation of the results.

spatial distribution of hits. A field of targets consisting of several rows would appear to be more sensitive to factors affecting the spatial distribution than would a pack in a single stack or even laid out in a single row. The research described here makes use of such an arrangement of the target cards in several rows.

Material and Methods

Instead of the ordinary ESP test cards or similar ones with a one-fifth probability basis, I used small white cards measuring 2 x 2 cm., each having the number *1* or *2* (a 1/2 probability basis) written upon its face. The numerals themselves were 12 mm. high. The

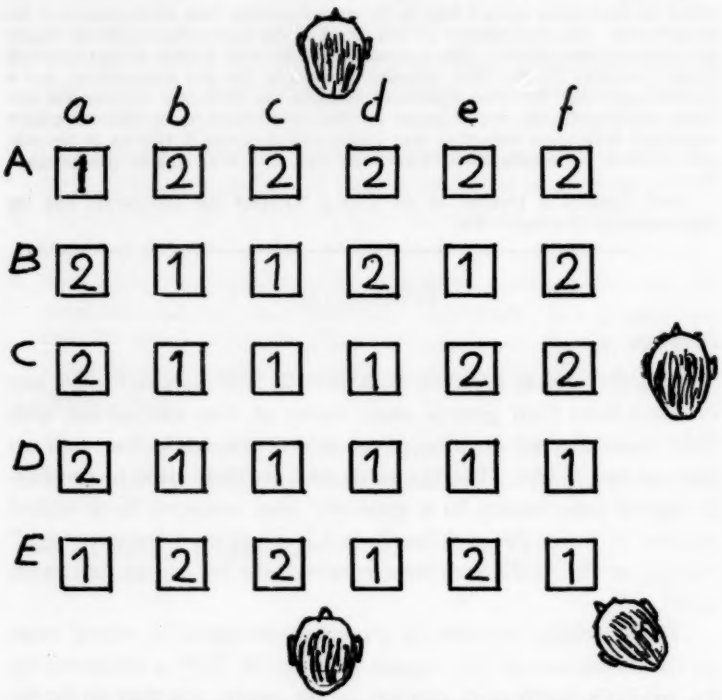

FIG. 1. Arrangement of target cards in rows and columns. (The lay-out of numbers in this case is only one of the many possible chance distributions.) The "heads" indicate four of the six positions relative to the targets which the subject was asked to take while he made his calls. The other two positions were underneath the cards and above the cards.

cards were put into individual dark paper envelopes which were thoroughly shuffled before each "run" and then put on a large tray in a rectangular arrangement consisting of five rows and six columns. (See Fig. 1.) Each envelope was separated from its neighbors by 5-6 cm.

The subjects (myself, my son, and my nephews and nieces) were tested individually. Each was invited to a session at what seemed a convenient time. I was careful to make the invitation friendly and informal; and if the subject showed any reluctance, the experiment was postponed until another day. The subjects did only one run of 30 cards at a session, and the speed of calling was chosen by the subjects themselves. They were instructed to name the numbers in the envelopes in ordinary reading order, starting with the top row and reading each row from left to right. They were not allowed to touch the envelopes and were encouraged not to look at them even when they might have been able to do so. In some cases, when the subjects had their backs to the cards or were situated under the table, or above it, the envelopes were completely out of sight. As the experimenter, I always made the record, and the subjects did not see the record until the end of the run, when the envelopes were opened and the cards were examined and recorded alongside the calls. The cards were shifted about among the envelopes many times at irregular intervals, so that there was no opportunity for the subjects to make associations between slight marks on the envelopes and the numbers inside.

Each run of 30 trials was made with the subject taking one of six positions relative to the 5 x 6 arrangement of the envelopes: facing directly; facing diagonally; facing directly away; at one side; with the targets *above* the subject; and with the targets *below* the subject. The first four positions are shown visually by the sketches of the heads shown in Fig. 1. Since this experiment was considered from the outset as being only exploratory, favorable psychological conditions seemed more important at this stage than considerations of nicety in statistical design. For this reason, the subjects were allowed to decide what position they wished to take in relation to the cards in each session. The only restriction imposed was that each of the child subjects was, by the time his part in the experiment

was finished, held to the same number of runs (six) for each position. In the preliminary tests which I made upon myself as subject, I did more trials and I also did not do the same number in all the positions (as will be seen when the results are presented).

Subjects

I worked first with myself as subject, then with eight children, 8 to 14 years of age, with whom I was on very good terms. The experiment lasted for three months during the summer of 1958. At the beginning I made some effort to classify the subjects in terms of their apparent ability by giving them several clairvoyance, precognition, and psychokinesis tests. Without analyzing the results precisely, I formed the impression that I myself was a middle-class subject (encouraging results in precognition and psychokinesis, poor ones in clairvoyance). My son Jean Pierre was by far the best in all tests. Next came my niece Antoinette M. and my nephew Laurence de C., then my nephew Phillippe M. (8 years old). Two other children, Louis M. and Geneviève M., were the worst, while a young girl, Eliane A., age 14, was very irregular. It is interesting to observe that Jean Pierre, Laurence de C., and Antoinette M. were gay and extroverted children, while Geneviève M. and Eliane A. were strongly introverted and showed periods of low spirits. I will not presume to classify myself psychologically!

I paid close attention to factors that might affect motivation and was careful to keep the experimental situation as familiar and informal as possible. The tests were not performed in the laboratory, but in the large living-room where the children were accustomed to playing. I also introduced a competitive factor by offering a prize each day to the subject with the best score. I spoke encouragingly after each run, whether successful or unsuccessful. It was with a view to reinforcing motivation that I had decided upon the use of the fifty-fifty basis of choice. Even the worst runs gave a number of hits that was not too far below 15 out of the 30 trials, and this was encouraging for the children. Also, the choice between two familiar digits, *1* and *2,* is a very simple one and does not bring into play so readily the conscious reflective thought processes. Altogether, these seemed like good conditions for the occurrence of ESP.

As van Busschbach (3) and many others have correctly pointed out, the relation between the subject and experimenter must be considered very carefully. My relations with my nephews and nieces are very good, and especially good with my son. The prizes (100 francs) were gained each day very easily, and that inspired the children to feel still more friendship for their uncle!

Aim of the Experiment

I can definitely state that the aim of the experiments was not primarily to obtain a total score with a high degree of statistical significance, but to see if the hits, regardless of the total number, would be randomly or evenly distributed over the rows and columns. This is, of course, what should happen if chance factors alone are present. As Gardner Murphy pointed out in a recent paper (1), the parapsychologists should not let themselves be hypnotized any longer to the point of looking only for statistical significance; they have to be more audacious. Sometimes a track might well be followed for a long way even if statistics do not give a *complete justification* for doing so *right from the beginning.*

Results

General

Although, as stated above, a significant result in terms of the total scores was not the primary objective of the experiment, it seems appropriate to follow the usual practice and to examine the data first from this point of view. The total number of trials made by all the subjects, including the experimenter, was 11,190, and the number of hits observed was 5567, a deviation of − 28 from mean chance expectation (M.C.E.). The CR of the total score is .53, which is not statistically significant.

When the results were broken down according to the individual scores, only two subjects were found to have noteworthy deviations from M.C.E, while the other six scored on the whole quite near to the expected number of hits. One of the two, J. P. C., had a positive score of 572 hits on 1080 trials (dev. = + 32), while the other, L. M., had a negative score of 511 hits on 1080 trials (dev. = − 29). A score at least as far from M.C.E. (in either direction)

as the former would be expected on a basis of random sampling about 5 times in 100; and the latter, about 8 times in 100. The individual results for the eight subjects are shown in Table 1. Each subject's total score is assessed in terms of the chi-square and the values have been added to get an over-all evaluation which takes account of the variation in scoring among individual subjects. This total chi-square is 8.41 which, with 8 degrees of freedom, has a P of only .4. Thus the total scores are well within the range of chance fluctuation.

Table 1

TOTAL RESULTS BY INDIVIDUAL SUBJECTS

Subject	No. of Trials	No. of Hits	% of Hits	χ^2 (1 d.f.)
R. C.	3630	1785	49.2	.99
P. M.	1080	548	50.7	.24
E. A.	1080	537	49.7	.03
A. M.	1080	544	50.4	.06
J. P. C.	1080	572	53.0	3.79
L. de C.	1080	534	49.4	.13
G. M.	1080	536	49.6	.06
L. M.	1080	511	47.4	3.11

$\Sigma\chi^2 = 8.41$ (8 d.f.)
$P = .4$

Hit Distribution Within the Columns and Rows

The results were divided into left-hand and right-hand halves of the target arrangement. Each half represented 5595 trials, so the expected number of hits in each group of three columns is 2797.5. A negative deviation of -82.5 hits was obtained in the three left-hand (a, b, c) columns, while a positive deviation of $+54.5$ hits was made on the three right-hand (d, e, f) columns. The CR of the difference is 2.59, with $P = .01$.

This tendency for the three right-hand columns to score higher than the others was found to occur consistently throughout the data. In a total of 57 runs in which different scores were obtained in these two halves, the higher score was made on the right 38 times and on the left only 19 times. A chi-square test applied to these figures indicates that only 12 times in 1000 would one expect a difference this large or larger from random sampling. In terms of individual subjects, one (G.M.) made the same score on each half,

six had more hits on the three right-hand columns, and only one (E.A.) had more hits on the three left-hand columns. Thus the effect of scoring higher on the right half of the target array was produced by most of the subjects and was spread over the great majority of sessions.

When the hits in the top two rows were compared with those in the bottom two rows, an insignificant top-bottom decline in scoring was found. The two top rows have a negative deviation of — 3, while the two bottom ones have a negative deviation of — 34.

To get a more detailed picture of the distribution of the hits over the target arrangement, the "quarter distribution" was examined. The results were found to be distributed as shown in Table 2. (The middle row, C, was omitted since it could not be assigned to any quarter of the 5 x 6 target arrangement.) From the table it can be seen that the largest departure from M.C.E. occurred in the lower left quarter, which has a deviation (— 51) of a magnitude that would on the average be equalled or exceeded only 3 times in 100 random samples from chance data.

Table 2

"QUARTER DISTRIBUTION" OF HITS WITHIN THE TARGET ARRANGEMENT

	Columns a, b, c	Columns d, e, f
Rows A, B.....	Hits = 1094 Dev. = −25	Hits = 1141 Dev. = +22
Rows D, E.....	Hits = 1068 Dev. = −51	Hits = 1136 Dev. = +17

Results for the Different Positions of the Subjects

A comparison of the total scores for the six different positions used shows that the highest *average* number of hits was obtained with the cards behind the subject's back (945 hits, or 45 more than M.C.E.), while the lowest scoring rate was obtained with the cards at the side of the head (851 hits, or 49 fewer than M.C.E.). Selected for separate evaluation and without any kind of correction for the background data from which they are taken, the former position has a P of .034 and the latter, .021. An adequate test of whether or

not position made any appreciable difference in the results may be made by means of a chi-square analysis applied to a 2 x 6 contingency table. This analysis, based upon the figures given in Table 3, shows that the position produced marginally significant differences in results ($\chi^2 = 13.69$, 5 d.f.; $P = .02$). These results indicate that position was probably a factor in the ESP performance, and they provide ample justification for giving some consideration to the reason for the differences.

Table 3

HITS AND MISSES IN THE TOTAL DATA IN RELATION TO THE POSITION OF THE SUBJECT

Position of Cards in Relation to Subject	No. of Hits	No. of Misses	No. of Trials	% of Hits
Facing	988	992	1980	49.9
At angle	1026	984	2010	51.0
Back toward	945	855	1800	52.5
At side	851	949	1800	47.3
Over	862	938	1800	47.9
Under	895	905	1800	49.7
Total	5567	5623	11190	49.7

$\chi^2 = 13.69$ (5 d.f.)
$P = .02$

DISCUSSION

The analysis of the results of the present experiment has revealed two general respects in which the data were non-homogeneous to a degree suggesting that something more than chance variation was involved. One is the difference between the first three and last three columns (an *increase* in the number of hits as the subjects read through the row of envelopes) with $P = .01$. The other is the difference among the results in terms of the positions of the subjects relative to the targets. In this respect the variation was of an order of magnitude that would be expected on a chance basis only 2 times in 100.

It is difficult, at this stage of what was undertaken as only an exploratory experiment, to advance hypotheses to account for the suggested nonrandomness in the data. But it is worth while to offer

some observations that appear to be relevant, even though they must be put forward tentatively.

One thing that must be considered is the general evidence for a psi-missing factor in the experiment: the negative deviation on the a, b, c columns was larger than the positive one on the d, e, f columns, and the second largest subject-score deviation was a negative one. As I pointed out at the beginning of the paper, my subjects appeared on the basis of preliminary tests to have different parapsychological capacities. I considered whether Schmeidler's work (2) might offer an explanation in terms of the attitudes of acceptance or rejection of the possibility of ESP; but there is another way to look at psi-missing for certain subjects. I think we would do well to keep in mind Warcollier's work on innate preference in barrage tests (4) and West and Fisk's paper (5) on the preference for certain hours in clock cards. One can never be sure there is not a preference for a symbol, even though the symbols selected for the test appear very homogeneous and psychologically equivalent. I analyzed the choices for the *1* and *2* numbers, and found an extremely strong preference for *2* in subjects L.M. and G.M. The higher scoring subjects (J.P.C., for one) are close to chance in their *1-2* choices in the total trials.

Since the subjects attempted to call the numbers by reading the rows from left to right, the difference between the a-b-c and d-e-f columns is comparable to the changes in scoring rate within the run as it has been frequently observed in ESP tests. In this experiment, however, an improvement in scoring rate was obtained as the subjects proceeded through the structured unit of calls (the 6-card row), while the trend that has been most commonly observed has been a decline in scoring through the column, or a U-shaped curve. An appropriate way of looking at the row in this case is to think of it as the *segment* of the run, and the most that can be said of the results at this time is that they suggest that a different segmental structure and a different order of calling might influence the position effects. The suggestion deserves further investigation.

Finally, what can be said regarding the suggestive difference among the subjects' positions in relation to the targets? It would be premature to offer the interpretation on the basis of the present data that the position of the subject's head influenced the scoring

on some yet-to-be-discovered physiological basis. In view of the absence of any evidence for any effect of position from previous researches, not only would it take much more evidence that this was the case than is provided by this experiment, but the differences that were found can more reasonably be attributed to psychological factors. No effort was made to keep the subjects from knowing which position was being used during each run, and I formed an impression as the testing proceeded regarding which condition was giving the best scores and which was the worst. It is not possible to say that this general observation on my part did not gradually build up an expectation regarding what kind of results each position would give; and any feeling of mine that one position was favorable and another unfavorable may well have been conveyed to the subjects in a subtle and even unconscious manner. Looking back upon this exploratory test, I can see that the use of the different positions seems to have served the purpose (an important one) of adding variety and interest to the experiment more than it did the aim of gathering data on a problem of basic research.

References

1. Murphy, G. Progress in parapsychology. *J. Parapsychol.*, 1958, **22**, 229-36.
2. Schmeidler, G. R., and McConnell, R. A. *Extrasensory Perception and Personality Patterns.* New Haven: Yale University Press, 1958.
3. van Busschbach, J. G. An investigation of ESP between teacher and pupil in American schools. *J. Parapsychol.*, 1956, **20**, 71-80.
4. Warcollier, R. Le problème des écarts negatifs en ESP. *Rev. Métapsychique*, July, 1955, 15-24.
5. West, D. J., and Fisk, G. W. A dual ESP experiment with clock cards. *J. Soc. psych. Res.*, 1953, **37**, 185-98.

Laboratoire D'Ethologie
Bures-sur-Yvette (S & O)
Paris, France

COMMENTS AND LETTERS

To the Editors:

This note is concerned with a friendly rebuttal of Mr. Dumont's recent letter,[1] as seen by a cyberneticist who has followed some of the work being done in parapsychology.

It seems to me that Cartesians have three principal complaints with parapsychologists. By a Cartesian I mean one who, for his conception of the world, has adopted, through value-judgment or indifference, a certain pair of basic metaphysical postulates from the philosophy of René Descartes. This pair of assumptions consists of a belief that (i) the whole is only some composition of its parts, or stated differently, there is no entirety apart from the relation of the parts; and (ii) knowledge is available only if a precise measurement can be made concerning some entity.

The first complaint concerns itself with experimental apparatus. Sometimes an apparatus seems to "violate" mechanistic principles where mechanistic principles seem to be useful. In other pieces of research some parapsychologists make use of conventional mathematics. It is at this point that some mathematicians believe that ESP is being used by some parapsychologists somewhere between their hypotheses and conclusions. Lastly, some data seem to be of such a nature that the "probability" of their being unconsciously altered is of the same order of magnitude as the "probability" of the event occurring.

To me, a scientist whose field is at least 100 years old and who deals only in the classical aspects of his field is a Cartesian. To me, a researcher whose field is concerned with such a contemporary model as the quantum theory of matter, the theory of political processes or stochastic processes—the double appearence of the word "process" being no coincidence—is certainly to be included among non-Cartesians. As far as I can gather a parapsychologist is basically a non-Cartesian. This is probably one of the principal reasons for his difficulty in communicating to "The Masses." However, if one is a non-Cartesian, this is neither a necessary nor a sufficient condition for making a useful contribution to a discipline. I con-

[1] *Journal of Parapsychology*, June, 1959.

sider a cyberneticist to be one who makes use of the best efforts of both the Cartesian and the non-Cartesian.

I am one of the first to admit that some of my fellow cyberneticists are *too* narrow-minded as to the potentialities of parapsychology. These cyberneticists believe, over-optimistically, that they are helping to prevent the uncorking of a farce. Their stand is tied up with the fact that Cartesian philosophy has, after three hundred years, seeped like disinfectant into the foundations of contemporary culture.

It is my faith as a philosopher, but not as an electrical scientist with my conditioned reflexes of a mechanist, that parapsychology will (i) provide some of the research for discovering, in addition to man's presently known senses, some for which he may have physiological capacity but which, as yet, he has not been taught to use; (ii) introduce a large class of new communication channels; and (iii) introduce a large class of new energy sources, each possibly as rich and useful as the existing electromagnetic source of energy.

In conclusion may I say that some of the items needed by parapsychologists are more papers criticizing parapsychology, more papers with experimental evidence, more papers on possible futures of parapsychology, and fewer papers by those people who would only criticize Cartesians for their cultural lag.

A. A. MULLIN

University of Illinois
Urbana, Illinois

To the Editors:

Mr. Mullin impresses me as a fair-minded critic who is insufficiently acquainted with the history and literature of psi research to form an accurate idea of the relationship between parapsychology and the structure of scientific thought in general. Since misconceptions on this point are by no means peculiar to Mr. Mullin, but rather characterize all our friends of the mathematico-mechanistic orthodoxy, it may be useful to attempt some clarification of the basic issues.

(1) Parapsychology is definable as an inquiry into phenomena that cannot be formulated in terms of space-time equations. ESP, for example, is perception, not only beyond the capacities of the

physical senses, but without the aid of electromagnetic radiations as these are understood by present-day science. The direction of the psi investigation itself demands the formulation of postulates of extraphysical factors and of hypotheses designed to test psi abilities to the limit. Parapsychologists will resort to these expedients without apology, inasmuch as the alternative would be to petrify the scientific system of thought in such a way as to give up from the outset any possibility of radically amplifying or changing it.

(2) The only relevance that I can assign to Cartesianism in the present discussion is the salutary shock that dogmatists might get from methodological doubt. Similarly, a dose of Cartesian dualism might have a therapeutic value for those who habitually resolve the mind-body relationship by suppressing the concept of mind. Such corrective measures, however, are clearly not what Mr. Mullin has in view, and I frankly do not see the pertinence of other aspects of Cartesianism to an examination of psi phenomena.

(3) In answer to the dictum that "knowledge is available only if a precise measurement can be made concerning some entity," one could hardly do better than to quote the chemist Standen: "If the idols of scientists were piled on top of one another in the manner of a totem pole, the topmost one would be a grinning fetish called Measurement." He goes on to point out that nothing is ever explained by showing how much larger or smaller it is than something else—which is all that metrology really is. The simple fact is that there are qualitative areas in nature which we can and must grapple with outside of dial readings and metrical data. In these areas unquantified concepts, far from being infertile, are likely to provide the only means of gaining vital insights. By the same token, the attempt to mathematize human personality and wipe out the distinction between mind and physical energy systems is precisely the error that has kept psychology from dealing adequately with the riches of the human subject. One is reminded of the pithy statement by Sherwood Taylor: "The materialistic hypothesis . . . is only made credible by discarding everything that does not agree with it. If I wish to support a theory that everything in a dining-room can be explained in terms of tables and chairs, my task will be much simplified if I describe knives and forks as illusions which will shortly prove to be aspects of tables and chairs, and if I take the precaution of de-

claring carpets and fire-irons not to be the subject of my investigation." In that case, however, one would arbitrarily shut out of one's experience whole ranges of important and valuable facts—which is exactly the stultifying performance put on by "scientific" materialists, despite their lack of support from modern theoretical physics.

(4) It must be emphasized that parapsychology, as an empirical inquiry, is under no obligation to concern itself with such philosophic issues as the dispute between monism and pluralism. More narrowly, the problem of the interaction between mind and body is not a road-block to keep us from looking into psi phenomena merely because no one at the moment can hazard anything but the vaguest speculations as to the nature of such interaction. All disciplines and sciences are abstractive. Whether they can ultimately be combined into a gigantic synthesis is a question that obviously cannot now be answered. Perhaps the complementarity principle forever bans a complete and unified reading of the cosmos. In any event, it is pointless to accuse parapsychology of not "fitting in" when the very meaning of "fitting in" has acquired a fatal ambiguity. As Bridgman has expressed it, the scientific method is "nothing more than doing one's damnedest with one's mind, *no holds barred* [italics mine]." Let our disingenuous obscurantists carefully reflect upon the meaning of this statement instead of inventing endless pretexts for shirking an open inquiry into an open universe.

T. DUMONT

P. O. Box 39
Madison Square Sta.
New York 10, N. Y.

To the Editors:

In connection with Carl von Klinckowstroem's translated article (in the March, 1959, number of the *Journal*) I would like to point out that the writer's rather severe criticism of the dowsing experiments conducted by the Royal Academy of Science of the Netherlands (p. 58) is not entirely justified.

These experiments should not be looked upon as an investigation into the whole problem of dowsing but have to be regarded as only testing the claims of a group of Dutch dowsers concerning the exist-

ence of mysterious *earthrays,* said to be the cause of cancer, various kinds of rheumatic diseases, etc. These dowsers asserted that the earthrays could only be detected by means of the divining rod (a thin metal wire was used for the purpose) and that if a man wanted to remain healthy the first thing to know was whether he was living on top of an earthray. This information was obtainable only through a dowser who, of course, had to be heavily paid to go over the consultant's house with his rod. If earthrays were detected (and they were invariably found in the bedroom under one's bed, or under the desk chair where one spends several hours a day working, but never in the W.C.!) one could escape growing a malignant cancer, etc. by buying a ray-screening apparatus, supplied by the dowser himself and costing up to one hundred dollars (the dowser's cost price usually was one dollar).

The whole thing soon became a complete racket and had such an influence on the general public and highly placed authorities that people *en masse* had their houses screened off in order not to run the risk of developing some serious illness. Governmental buildings were having anti-earthray apparatus fitted at great cost, while even the Royal Family had dowsers over to run the damaging rays to earth and protect themselves by having the various apparatuses installed. Farmers, cattle-breeders, bulb-growers, etc. were paying out large sums of money because they were made to believe that crop failures, sickness of live-stock, etc. were caused by the disease-producing rays. The Dutch public began to panic, threatened on all sides by this mysterious but deadly earthray, and it was then that the Dutch Government requested the Royal Academy to appoint a Committee to investigate the claims of the Dutch dowsers, i.e., to detect by means of their rods physical rays or agencies causing the diseases. (The dowsers even claimed that they could indicate what kind of disease the rays produced by counting the turns of their rods!) The first thing to do—and rightly so—was to see whether the dowser and his rod did react to some objective reality. The Committee's experiments, I believe, were quite up to the mark and, as could be foreseen, gave nothing but negative results. The earthrays, in the Committee's opinion, existed only in the dowser's imagination.

It is indeed interesting to note that after the publication of the Committee's negative results, the earthray epidemic in Holland died down completely and nowadays not a single person talks any more about the cancer-producing rod-sensitive earthrays. The five thousand or so Dutch dowsers, however, are crying their eyes out for having been robbed of a very lucrative business.

G. ZORAB

Verdistraat 32
The Hague, The Netherlands

BOOK REVIEWS

THE SIXTH SENSE. AN INQUIRY INTO EXTRA-SENSORY PERCEPTION. By Rosalind Heywood. London: Chatto and Windus, 1959. 224 pp.

There have been many summaries of psychical research by persons who were not familiar with the field. The result has been somewhat warped interpretations of the evidence. This short history is not of that sort. In a rather curtailed space, Mrs. Heywood has outlined very carefully the significance of the inquiry into extrasensory perception from its academic beginnings in the Ghost Society of Cambridge University one hundred years ago.

Perhaps it was Professor Henry Sidgwick's interest in psi phenomena which began in this small group that led to the founding of the Society for Psychical Research in 1882. Mrs. Heywood traces with considerable familiarity the progress of the research from those beginnings, with sharp discernment as to the highways encountered and the by-ways to be avoided. She describes in some detail the early experiments of Sidgwick, Frederic Myers, and Edmund Gurney in quantitative and qualitative approaches to investigation. She describes the development in 1885 when Oliver Lodge applied probability mathematics to the results of an experiment in extrasensory perception.

One of the most intriguing parts of this book is the clean-cut interpretation and evaluation of the cross-correspondence experiments of the S.P.R. group between 1901 and the mid-twenties. Of unusual interest is the chapter on the Willett scripts, published only in part by the S.P.R. and analyzed by Gerald, Earl of Balfour, who conducted this inquiry. Inasmuch as Mrs. Willett is dead, Mrs. Heywood suggests that the unpublished portions of the scripts can now be printed. "Mrs. Willett" was a pseudonym for Mrs. Coombe-Tennant, a most unusual sensitive.

It is to be regretted that the latter parts of this book, which deal with the quantitative experimental work in extrasensory perception and psi phenomena by J. B. Rhine and the Duke University group, as well as those of G. N. M. Tyrrell, S. G. Soal, Whately Carington,

and Robert H. Thouless, are too compressed to do full justice to this important line of investigation.

Mrs. Heywood writes with clarity and with brilliant insight into some of the obscure factors which are active in the field of psi. Her concluding chapter, "The Future," outlines possible new approaches for investigation and accents the need for immediate studies in special areas where psi-inhibiting skepticism is on the increase.

This is not just another historical summary of investigations into ESP. Mrs. Heywood's ability to choose cases of exceptional merit and explain them simply and adequately illumines the book and makes it very worthwhile reading.

EDMOND P. GIBSON

1221 Philadelphia Ave., S.E.
Grand Rapids 6, Michigan

WATER WITCHING U.S.A. By Evon Z. Vogt and Ray Hyman. Chicago: Univ. of Chicago Press, 1959. 248 pp. $4.95.

This book, written by an anthropologist and a psychologist, presents a survey of water dowsing or, as its authors prefer to call it, water witching, in rural America. In reading this book it would be well to keep clearly in view the distinction between an erroneous belief, rationally arrived at, and a superstition. For in this book dowsing is arbitrarily treated as sheer superstition or as a "magical ritual" based on folklore. To those who unqualifiedly accept this thesis the book is an able and interesting sociological and psychological study of an apparently irrepressible folk phenomenon. To those who do not fully accept this thesis the book fails as a scientific study due to an unscientific display of bias and prejudice.

Like a criminal at the bar, dowsing, before being convicted, should have the benefit of a reasonable doubt. That it ought not to be condemned in cavalier fashion is affirmed by those scientists of the first rank who have personally investigated it. In this connection Sir J. J. Thomson wrote in his autobiography:[1] "The [dowsing] rod may be a convenient indicator of a psychological effect. . . . Although I think that most of the people who have paid any attention to the subject believe in the reality of dowsing, there is no agreement about its cause."

[1] *Recollections and Reflections.* New York: The Macmillan Co., 1937.

There are two alternatives under consideration. If we assume that dowsing contains a valid substratum of truth, it must nevertheless be admitted that around the practice there has grown up a large body of superstition and folklore, like that which surrounded astronomy, chemistry, and medicine in earlier times. And this fact justifies a scientific study like the one reported in this book. If, on the other hand, it should eventually turn out that there is no rational basis whatever for dowsing, there is still no justification for the general tone of this book. It requires little courage to stand entrenched in "orthodox" science and to ridicule those scientists who have fearlessly faced the possibility that even the weirdest phenomena may contain the germs of future intellectual progress. One thinks of the story of Mesmer, much publicized in recent decades. Buried in his fantastic theory of "animal magnetism" was the phenomenon of hypnotism; yet not one of the leading scientists of the day who unanimously condemned Mesmerism took the trouble to investigate his claims experimentally. As a result, hypnotism was buried in a fog of superstition until exhumed by Charcot, Janet, and Freud three quarters of a century later.

This book, as stated, carries throughout its length an overtone of prejudice that is sadly at variance with the true spirit of science. Prejudice, like injustice, is bad in itself and is never justified by the end results. It also leads almost inevitably to various errors, such as unconscious distortion of evidence to favor a preferred thesis. The reader will discover many examples in the book. Scientists of outstanding reputation, who have personally investigated dowsing and have ventured to conclude that it merits further study, are treated with contempt. The meticulously honest reporting of the late Kenneth Roberts is condemned or ignored. The problem of dowsing is approached statistically, as if it were a mass phenomenon and not a matter of individual talent.

The first and perhaps the most conspicuous example of bias is the authors' use of the term "water witching" for dowsing. This is a colloquialism which the authors found to be more frequently used in rural areas than any other term, and on this basis they justify its use. But these authors, who are highly competent in the social sciences, must know that this usage is a semantic device tending to discredit the practice and to spread prejudice. If further evidence

is needed it is only necessary to cite such chapter headings as "Who's Who in Witching" and "For Whom the Witch Tells." This sort of thing is completely out of place in a scientific book.

There are some significant omissions from the book. Possibly the authors did not discover cases of "dowsing" without the use of a forked stick or other device. Yet several such cases are recorded in the current literature. A remarkable one was that of the late J. Raoul Desrosier (see Kenneth Roberts' *The Seventh Sense*), who claimed to be able to find water by curious sensations in his chest, resembling the pains of pleurisy. These cases require a modification in the definition of the word dowser. But perhaps the strangest omission of all is this: Although the alleged phenomena of dowsing are clearly in the closest relationship to the alleged phenomena of clairvoyance and telepathy, or ESP, the authors avoid a discussion of these phenomena. The widely known experimental work of Dr. J. B. Rhine and his associates at Duke University, carried out to meet exacting standards and verified in other laboratories, is ignored. This can only be taken to mean that the authors do not concede even the possibility that ESP exists. This position is in marked contrast to that taken by some other scientists who are quite as eager to reject such beliefs. Referring to the phenomena of ESP in connection with the theory of computing machines, the late British logician A. M. Turing said: "These disturbing phenomena seem to deny all our usual scientific ideas. How we should like to discredit them! Unfortunately, the statistical evidence, at least for telepathy, is overwhelming."[2]

To avoid misinterpretation of what has been said here, it should be understood that no exception is taken to the fact that these authors do not accept the reality of dowsing, for they have the right to interpret the evidence in this way. What is unfortunate, and what detracts greatly from the value of their book, is their uncompromising elimination of the *possibility* that there exist at least a small number of dowsers with the unexplainable ability to locate water in an extrasensory manner.

The constructive sections of this book report on the results obtained from a questionnaire mailed to five hundred carefully selected

[2] "Can a Machine Think?" in *The World of Mathematics*, Vol. 4. New York: Simon and Schuster, 1956. See p. 2116.

county agricultural extension agents. The principal conclusion is that dowsers, as a class, are in general conformity with their communities. A further conclusion is that "in every case known to us, the diviner is an honest man of recognized integrity. He practices his art out of a sincere desire to help his neighbor."

For the rest, the book is devoted to answering four questions: What is water witching? Does it work? What makes the rod move? Why does water witching continue to be practiced in the United States? Much of the material is prepared for the general reader. The classical theory of unconscious muscular motions to account for the gyrations of the rod is developed at length, with excursions into related fields. Practically all authorities are in agreement here.

Due to restrictions of space this review does not perhaps do justice to the better features of this book, and for such shortcomings the reviewer wishes to apologize. He also wishes to explain that in his view the most important issue involved is the intrusion into scientific discourse of "scientific authoritarianism" and prejudice. These twin diseases could destroy all of science. That the authors are not unaware of this danger is shown by their reference (page 215) to "the precautions that are necessary to pursue truth unfettered by our prejudices."

HORACE C. LEVINSON

Kennebunk, Maine

PARAPSYCHOLOGIE. By G. Zorab. S'Gravenhage: Universiteit voor Zelfstudie, 1959. 368 pp.

The author of this book is well known in the world of parapsychology. After first working in several other professions, he has for some time dedicated himself to the study of this field. Zorab has published a great number of books and articles on the subject in Dutch, English, and American periodicals, and he has also directed a number of investigations in ESP. It is the purpose of the book (written at the request of the editor of the University of Self Study at the Hague, Netherlands) to give the well-informed layman a review of the different areas of parapsychology. The bibliography is so detailed (230 titles) that anyone who wants to occupy himself

extensively with one or more aspects of the field can easily orient himself.

The preface (Part I) gives a short review of the different phenomena included in parapsychology, classifies them, and gives the origin of some of the terms.

The author divides the phenomena into parapsychical (Part II) and paraphysical (Part III). A large part of these sections is given over to the description of phenomena, investigations, and tests. The manifold citations of remarkable evidence allow the book to be read with ease and pleasure, but at the same time, they make it a little too anecdotal in some parts. Besides, many of the examples date back very far (as far as the fourth century) and are subject to all the objections of citing old historical cases. If the author had used only examples from the twentieth century (and there are enough to be found) the contents of the book would have been more factual and more stimulating.

Fortunately, the author's own critical and analytical attitude forms a valuable counterbalance and offsets two important forms of reaction from which parapsychology is still suffering (although in a clearly diminishing measure): on the one hand, a strong, emotionally charged defense of it which is without objective, critical discernment; on the other hand, an equally strong, emotionally charged *a priori* rejection of the whole problem without objective, critical, scientific examination of the facts. In his book, Zorab tries to be objective and realistic and to get to the heart of the matter, and in most cases he succeeds very well.

Part II divides parapsychical phenomena into ESP, precognition, sensory and motor automatisms, and psychometry. After some historical cases, the author goes on to the statistical experiments. The explanation of the binomial method is rather short and does not go into detail. Is it a misprint that in the explanation of the formula $\sqrt{npq}$, p is called the number of hits instead of the probability of a hit on each trial, and vice versa for q? No other methods of statistical computation are described. It is possible that this is due to the author's fear of making the book too technical. The description of the tests by van Busschbach contains a disturbing error: the number of hits on the +1 cards is called significantly high (CR= 3.3) whereas the result was significantly low (CR= −3.3).

Extensive treatment is given to the use of the words "psychometry," "psychoscopy," "paragnosy," and "ESP" as well as to the confusion caused by the different interpretations of these words. The author warns against this confusion, but he himself does not avoid it; for example, where he calls Croiset a psychometrist at the beginning of a short sentence and a paragnost at the end.

Part III, containing the paraphysical phenomena, is divided into physical effects, an extensive treatment of the most important mediums (the Schneiders, Eusapia Palladino, etc.); PK; paranormal healing; apparitions of the living, dying, and dead; ghosts; and experimental investigation of the question of survival. It begins with the statement that parapsychology owes its origin partly to modern spiritualism or spiritism, which in turn grew out of the mediumistic phenomenon of rappings. This statement will meet with opposition, we suppose. Here also the author tries to be as objective as possible. He himself introduces the discussion by indicating that this is not a field for exact scientific research only, but also for philosophical interpretations of life and of man. The different subtitles of this part touch upon many questionable points and indicate that belief in certain explanations is often a far different thing than actual knowledge.

After an extensive description of the achievements of mediums, and of people who might pass for mediums, and a serious warning against fraud and lack of adequate control, the book ends with a critical treatment of the very recent proposal by Dr. Karlis Osis about the investigation of the question of individual survival and with the statement that the increasing proofs of the reality of precognition seem to undermine the traditional opinion of the universe with its images of time, space, and causality to such a degree that the spiritistic ideas about survival as well as the scientific methods of research undoubtedly will feel their influence.

Like a lighthouse, Zorab's book swings its rays over a wide sea of problems and questions. It gives the reader a general orientation, and if the reader intends to land on the beach of scientific research, this book will pilot him to a landing on his own special sector.

J. G. VAN BUSSCHBACH

Kleverlaan 151
Haarlem, The Netherlands

PARAPSYCHOLOGICAL ABSTRACTS

SARA R. FEATHER

[This section consists of abstracts of reports published in other journals and of unpublished reports filed at the Parapsychology Laboratory at Duke University. Readers may obtain unabridged copies (Thermo-Fax) of the unpublished reports by writing to the *Journal of Parapsychology,* College Station, Durham, N. C. The price is ten cents per page.]

Ducasse, C. J. How good is the evidence for survival after death? *J. Amer. Soc. psych. Res.,* 1959, **53**, 90-105.

As his basic premise, Ducasse asserts that survival of consciousness after death remains theoretically and empirically possible because the supposition of it is not internally contradictory nor incompatible with anything definitely known. The question of survival for how long and of just what is closely examined, and possible kinds of "life" after death are considered. Empirical evidence of survival, if any, will have to be found, Ducasse maintains, among certain kinds of ostensibly parapsychical occurrences, chief of those being apparitions, "out-of-the-body" experiences, "possessions," and communications obtained through mediums, notably the "cross-correspondences." Ducasse outlines the factors upon which he thinks the final answer will depend in deciding between the survival hypothesis and the "only hypothesis alternative" to survival, the telepathy-clairvoyance explanation. He concludes that the balance of evidence so far obtained favors survival, not of mere memories, but of "the most significant capacities of the human mind, and of continuing exercise of them after death."

Fahler, J. Exploratory "scaled" PK placement tests with nine college students with and without distance. *J. Amer. Soc. psych. Res.,* 1959, **53**, 106-13.

Confirmation of the PK placement hypothesis is reported in an experiment in which degree of success was measured by the distance in centimeters that the dice cubes rolled in the correct direction (A- or B-side) on the PK table. Nine especially selected college students

completed five sets of releases (50 throws of 6 dice) under each of two conditions: a "no distance" condition in which the subject was in the same room as the apparatus, and a "long distance" condition when the subject was in another room, placing him 27 yards from the apparatus. On the basis of their performance in an initial pilot study, the direction (whether positive or negative) of the subject's scoring in the "no distance" releases was successfully predicted (P=.0015). Both groups, the four subjects who were predicted to score negatively and the five predicted to score positively, obtained significant results in the predicted directions. Out of the total 45 sets in the "long distance" condition, 33 were positive (P=.0002). The author notes that strong feelings of sympathy and empathy existed between himself and the subjects and that this may have been a favorable factor in the experiments.

Lancaster, J. B. A GESP experiment with a dual (color-symbol) target. (Unpublished manuscript.)

Although there were no clear-cut significant findings, some marginally significant results were obtained in a two-part, long-distance GESP experiment involving a total of 34 subjects who guessed the order of dual targets that consisted of combinations of one of five colors with one of five symbols. A chi-square analysis of the first series showed results (χ^2=2.51, 1 d.f.) in the predicted direction, in line with the author's previous findings that subjects who had scored high or low on one aspect of the target tended to score high or low, respectively, on the other. In the second series, which tested the 10 highest scores of the previous series, the P value of this same breakdown in the two-by-two table was .022 when corrected for selection.

THE PARAPSYCHOLOGICAL ASSOCIATION

The Second Annual Convention[1]

The second annual convention of the Parapsychological Association was held at the Hotel Martinique, New York City, September 10-12, 1959. The program was as follows:

Program

Thursday, September 10

I

9:00 Registration

9:30 Convention Call to Order...Dr. G. R. Schmeidler, President
Secretary's Report.....Mr. W. G. Roll, Secretary
Treasurer's Report.....Dr. R. A. McConnell, Treasurer

II

Chairman: Dr. J. L. Woodruff (City College, New York)

10:30 "Awareness of Hitting in Precognition Experiments with Hypnotized Subjects" by Mr. Jarl Fahler (University of Helsingfors, Finland)[2]

Two subjects were specially selected from a previous experiment to take part in a precognition test in which the targets were to be chosen one day after the calls were made. Before the experiment, each subject was hypnotized, and it was suggested to him that when he took the test, he would have a special impression of "correctness" or "difference" for some calls and that he was to have these calls marked on his record sheet. Both subjects got insignificant negative deviations on their total guesses for the target. On the marked ones, both the subjects scored positively and their combined scores were highly significant. The success on those that were marked was interpreted as meaning either that the calling in these instances was accompanied by a conscious experience of success or something approaching this; or that the subject was good at guessing by precognitive ESP which of his calls were going to be right the next day.

[1] This summary was prepared by Dorothy H. Pope.
[2] Presented by Dr. Karlis Osis.

11:00 "Diametric Precognition" by Dr. C. B. Nash (St. Joseph's College)

The general objective of an experiment in precognition was to determine whether the subjects could hit future targets selected by a distinctly two-step procedure. The targets were the digits 0-9. Two days to two weeks after the subject wrote his calls on a record sheet, a randomly selected digit was transcribed next to each one. Each of these digits referred to the number on an envelope containing a blank card. One and a half to two years after the subject made his calls, the target digits were selected and written on the blank cards. As additional tasks, the subject had been asked to estimate whether his calls would be above average, average, or below average, and also to check those he felt especially sure of. The scoring on the actual targets and on the estimates of scoring level were not above chance. Analysis of the distribution of success among the checked and unchecked calls, however, showed significant differences and these results were interpreted as indicating that psi had functioned diametrically in two acts of precognition separated by an interval of one and a half to two years.

11:30 "Rorschach Records and ESP Scores" by Dr. G. R. Schmeidler (City College, New York)

Group ESP tests and Rorschach tests were administered to subjects classed as sheep or goats. A test group of 1004 subjects confirmed the hypothesis that sheep with good social adjustment tend to have higher ESP scores than other sheep and that goats with good social adjustment tend to have lower ESP scores than other goats. Further detailed analyses of the Rorschachs imply that freely responsive subjects tend to have high ESP scores whether they are sheep or goats and that markedly restrained or inhibited subjects tend to have low ESP scores whether they are sheep or goats. A tentative interpretation is that the feeling of free acceptance of the over-all situation is likely to be associated with high ESP scores and that feelings of constraint and withdrawal tend to be associated with chance ESP scores or with psi-missing. Presumably, well adjusted sheep are more likely to respond with a free acceptance of the ESP situation than goats or poorly adjusted sheep.

III

Chairman: Dr. R. A. McConnell (University of Pittsburgh)

1:30 "Statistics and Model Building in Parapsychology" by Dr. M. E. Turner (Medical School of Virginia)

An underlying mathematical model is assumed for all modes of statistical inference, whether this mode of inference be that of (1) estimation; (2) tests of significance (or tests of hypotheses); or (3) confidence limit determination. The model is a symbolic description of some aspect of a system, usually an algebraic description, chosen ideally in such a manner that all acceptable hypotheses (acceptable descriptions) are special cases of the model, and such that all unacceptable hypotheses are not special cases of the model. Thus all a priori information about the system is contained in the model. All unknown information is represented by symbolic constants, termed parameters, appearing in the model. Any model not closely approaching the ideal one is inadequate in the sense that estimates of meaningful parameters are not forthcoming, tests of significance are insensitive, and confidence limits will be too broad. When the model is adequate, great facility in posing questions to be answered by experiment or observation is gained. For any conceivable question there is a corresponding statement involving only the parameters of the model. Appropriate methods of statistical inference will then be dictated. Some examples of simple models in parapsychology were discussed. These include a model for describing the change in scoring level in a series of runs, and a model for describing the conceivable relations between scoring level and distance.

2:00 "Position Effects in PK Placement Experiments" by Mr. H. Forwald (Ludvika, Sweden)

Previous reports have dealt with a long series of PK placement tests in which a group of electrically released dice were "willed" by the subject to go to one or the other side of a divided dice table. These data have already been extensively studied from the point of view of the energetics involved in the transverse displacement of the dice. The latest research studies the longitudinal displacement.

A similarity has been found in the aberrations of the dice in the two directions, suggesting that the subject's mental efforts work to some extent in the longitudinal direction too. The aberrations here are smaller, however.

2:40 "Mechanized and Integrated Three-Tier Placement PK" by Mr. W. E. Cox (Southern Pines, N. C.)

Improvements have been made on the original "three-tier" placement apparatus (see *J. Parapsychol.*, March, 1959) and a further series of placement tests has been carried out. Dice are released into the top tier of the apparatus. The subject tries by PK to make them drop on to the target half of the top tier, which allows them to drop to the second tier. If they fall on the target side of this tier, they are free to fall to the bottom tier, the surface of which is a checkerboard with alternate squares as targets. It was hypothesized that the proportion of dice which scored hits would be greatest in the bottom tier. The results, though not significant, were in the expected direction and tend to confirm the previous findings. Photographic records of the dice as they came to rest on the bottom tier show placement effects that are interpreted as lending further support to the PK hypothesis.

3:10 "Experiments Concerning ESP and Pharmacology" by Dr. C. Wilson (University of Liverpool, England)

In order to apply the methods or techniques of other scientific disciplines to parapsychology, the investigator must make the assumption that the observations made by parapsychologists are manifestations of a property inherent in man. The hypothetical center where psi phenomena originate would conceivably lie in the reticular activating system where those peripheral impulses which enter conscious awareness are filtered off from those which remain subconcious. Any drugs which affect the function of this part of the central nervous system might therefore be expected to affect ESP ability. Two experiments were carried out in which the subjects were first given a control ESP card test. Then they were tested in a series of sessions in which they were given an unidentified drug, followed by another ESP test. The drugs used were: dexampheta-

mine, quinalbarbitone, caffeine, a tranquilizer, and lactose as a control. Statistically significant evidence for the action of drugs on ESP ability was not obtained, but the experimental results provide suggestive evidence that more extensive testing or more active drugs may provide significant results.

3:40 "An Experiment with Targets That Differ in Degree of Similarity" by Dr. Catherine Nash (St. Joseph's College)

Eight animal symbols stamped on blank playing cards constituted the targets in this clairvoyance test. As an aid to the subject, an illustration sheet bearing the same symbols (in two rows of four symbols each) was placed before the subject as he took his test. Although the experiment was designed primarily to find out whether the subject would tend to confuse symbols that were similar more than he would those which were dissimilar, no extrachance "consistent missing" effects were found. The most interesting results (which were not expected) were related to the illustration sheet. On those targets which happened to be the symbols in the four corners of the sheet, there was a marginally significant positive deviation in scoring. The targets which were shown in the inner positions on the illustration gave scores close to chance. Correlations of the ESP scoring level with the personality traits measured by the Minnesota Multiphasic Personality Inventory yielded a correlation between ESP scoring level and social introversion which is significant at the .006 level of probability (not corrected for selection).

7:00 Convention Dinner Address: "Where Do We Go Next in Parapsychology?" by Dr. R. H. Thouless (Cambridge University England)

One way of evaluating the future of parapsychology is to take a lead from the criticisms that are made of the field. Current skepticism toward parapsychology will not be reduced by more evidence of the kind we already have; we need experiments of a different sort. The present-day criticisms suggest the following new paths for exploration: (1) We should provide a pool of successful scorers to be made available to skeptical scientists who wish to experiment under their own conditions. (2) We should try to discover methods of

getting reliable ESP results. Self-experimentation appears to be the most helpful approach. (3) A great deal depends on getting psi to work. The repeated-guessing technique offers the best prospect of bringing ESP out of the laboratory and applying it to the practical task of acquiring and communicating information.

Friday, September 11

I

Chairman: Dr. M. Scriven (Swarthmore College)

9:00 "On the Question of Control Over ESP" by Dr. J. G. Pratt (Parapsychology Laboratory, Duke University)

The first 22 volumes of the *Journal of Parapsychology* were reviewed in order to list individual high-scoring ESP subjects. Thirty-four persons were found who were considered as belonging in this category. Their results, the nearest approach to voluntary control over ESP yet achieved, were examined in relation to a single question: Was the subject's familiarity with the environment in which he was tested a factor in his ESP success? For three of the high-scoring subjects listed, no information relevant to this question was available. Among the remaining 31, 26 had been tested either in their own homes or in comparably familiar surroundings. Two earlier investigators (Tyrrell and Stuart) discussed the importance of testing subjects in their own environment, and the present survey of the experimental evidence strongly underscores their suggestion as an important one, even though it has been largely overlooked by research workers in parapsychology.

9:30 "The Repeated-Guessing Technique" by Dr. R. H. Thouless (Cambridge University, England)

Repeated-guessing on an individualized pack has been used in the hope of increasing the significance of weak scores. Actually, significance cannot be increased by this method. Its usefulness is with high-scoring results to bridge the gap between mere probabilities and real information. This gap must be bridged if ESP is to leave the laboratory and do real work. Concentration by the repeated-

guessing technique is illustrated by Dr. Soal's experiment with Mrs. Stewart in which she was guessing playing cards.

10:00 "Towards a New Hypothesis" by Mr. K. Ramakrishna Rao (Andhra State University, India)

Many hypotheses that are put forward to explain psi start with the assumption that psi is nonphysical, nonsensory. Investigations involving physical targets and concerned with the physical, physiological, and psychological conditions affecting psi are inconsistent with the claim that psi is nonphysical. These and other logical difficulties raised by the physical-nonphysical dichotomy become irrelevant if one considers that psi may be a kind of simple imagination generated at deep levels of the unconscious and stimulated to activity by our nervous energy (and possibly by volition) via some as-yet-unrecognized route which then leads the way back up to the conscious level. Such a hypothesis does not involve the paradoxical dichotomies of physical and nonphysical; mental and material; sensory and nonsensory.

10:40 "Towards an Explanatory Model for Psi Phenomena" by Mr. W. G. Roll (Parapsychology Laboratory, Duke University)

Studies of the ESP process fall into two categories; one concerns the nature of the stimulus and the other, the nature of the subject's response. Recent research interest has centered on the latter area, and analyses of spontaneous cases as well as experiments involving psychological tests and attitude questionnaires have shed much light on the factors that determine which form the response will take. However, very little attention has been paid to the stimulus that evokes psi, since it became apparent that psi is probably not dependent on known physical laws. On the basis of a reexamination, mainly of the older experiments of Hodgson, Soal, Osty, and others with such subjects as Mrs. Piper and Mrs. Cooper, it was suggested that there may be external conditions that do not fall within the scope of present-day physics on which the transmission of ESP stimuli may depend. The theory, which involves the postulate of new constituents of material objects with new causal properties, is also intended to provide a unified description of the various types of psi phenomena, such as ESP, PK, and precognition.

11:10 "Clairvoyance Scores of Sheep versus Goats When Subjects's Attitude Toward the Experimenter and the Purpose of the Experiment Are Manipulated" by Mr. Sherwood Waldron, Jr. (Yale University)

Five groups of college students were tested for clairvoyance to see if the subject's reaction to the experimenter was not an important variable in determining whether "sheep" would tend to score higher than "goats" as found repeatedly under certain conditions by Schmeidler. It was hypothesized that the subject's scoring might be influenced by his reaction to the experimenter and that this in turn might be affected by whether he thought the experimenter agreed with him or not in his beliefs about ESP. Therefore, the groups were told different things: the first and fifth were told nothing about the experimenter's beliefs; one group was told he did not believe in ESP; one was told he did believe in ESP; and one, used as a control, was led to believe they were being tested for subliminal vision. Two out of four of the interactions calculated to test the hypothesis were significant. Thus the results tended to bear out the hypothesis that the sheep-goat difference is determined more by differences in interpersonal relations between subject and experimenter than by any direct effect of belief on scoring.

11:40 "Associative Aspects of Psi Phenomena" by Dr. J. H. Rush (Boulder, Colorado)

The associative aspects of psi, including those that have been noted in detail by Jung, Carington, and Ehrenwald, suggest that psi phenomena relate to the free-associative processes of the unconscious in a way that is equivalent to the relation between the causally determined phenomena of physical science and the rational, logically determinate processes of the conscious mind. Some data further suggest that sensorimotor (SM) and psi processes tend to exclude or complement each other, the one mode of cognition or expression being enhanced insofar as the other is in abeyance. This concept implies that both SM and psi components probably are jointly involved in ordinary behavior, and that experimental situations that permit this familiar interaction should be more favorable for obtaining psi manifestations than are those in which the SM

component is completely excluded. The experiment of Brugmans and Heymans can be interpreted as a fortunate instance of such interaction between psi and SM functioning.

II

Chairman: Dr. Gardner Murphy (Menninger Foundation)

1:30 "The 'Linkage' Effect in ESP" by Dr. B. C. Kirby (San Diego State College)

This experiment was a test of the transmissibility of information extrasensorially through 0, 1, 2, or 3 "links" or "removes" in the information chain between subject and agent. Ordinarily the parapsychologist asks if information may travel between subjects A and B, with both A and B and their locations being personally known to each other (the zero-remove situation). This research asks if A can communicate parapsychically with a third subject, C, whom he does not know (and whose location he does not know) but who *is* known to B (the one-remove situation); or with subject D (known to C but not to B or A); or even with E, three steps removed. In 21,400 trials, distributed among 0, 1, 2, and 3 removes, results highly suggestive of ESP occurred clearly at the zero remove, not at all at two or three removes, and inconsistently at one remove.

2:00 "The Distance Problem" by Dr. Karlis Osis (Parapsychology Foundation, N. Y.)

A survey was made of published distance experiments in which targets with 1/5 probability of hitting were used. The data were grouped in distance classes: 100 yards to 1 mile; 1 mile to 1000 miles; and so on at 1000-mile intervals. Scoring level declined with each distance group approximately corresponding to the inverse square law. However, the trend might be explainable by a combination of subject and attitude variables. The authors' own experiments that were reported were not subject to interpretation in terms of those factors. The over-all trend of results indicate that distance might influence ESP scoring, but definite conclusions have not yet been reached.

2:30 "A Repetition of the Anderson-White Investigations of Teacher-Pupil Attitudes and Clairvoyance Test Results" by Mr. A. DeGuisne and Mr. G. Goldstone (MacMurray College)

[This report was published in the September, 1959, number of the *Journal of Parapsychology,* pp. 196-214.]

3:10 "Teacher-Pupil Attitudes and Clairvoyance Test Scores in College and Elementary School Classes" by Mr. M. Rilling and Mr. J. Adams (Oberlin College)

In clairvoyance experiments with elementary-school and high-school pupils as subjects, Anderson and White have reported that the level of scoring is associated with differences in teacher-pupil attitudes. The experiment described here was an attempt to repeat the Anderson-White tests with a group of fourth- and fifth-grade pupils and a group of college students as subjects. The ESP results were not significant for either group. The elementary-school data revealed directional similarities with the Anderson-White research in parts of the data. In the college results, when the professor was positive toward ESP, his class tended to score above chance; and when he was negative toward ESP, the results were below chance. Suggestions for methodological improvements in classroom experiments were noted.

3:40 "A Two-Year Program of Tests for Clairvoyance and Precognition with a Class of Public School Pupils" by Miss Margaret Anderson (University of Pittsburgh) and Mrs. Elsie Gregory (Hawthorne Elementary School, Wheaton, Illinois)

[This report was published in the September, 1959, number of the *Journal of Parapsychology,* pp. 149-77.]

4:30 Presidential Address by Dr. Gertrude Schmeidler (City College, New York)

[The full text of this address appears on pp. 238-250 of this number of the *Journal.*]

Saturday, September 12

I

Chairman: Dr. C. J. Ducasse (Brown University)

9:00 "Parapsychology Course on the Undergraduate Level" by Dr. J. Freeman (Wayland College)

In the fall of 1957 the Psychology Department of Wayland College, Plainview, Texas, first offered a three-hour credit course entitled "Introduction to Extrasensory Perception." In addition to the textbook (the Rhine-Pratt book *Parapsychology*), the course included outside reading, discussion, group experiments, etc. Each student was asked to design and conduct an original experiment. Those students whose projects were sufficiently encouraging were allowed to continue them in a course in Experimental Psychology the following semester. The plan now in effect requires the student to take the Experimental Psychology course if he is to receive credit for the course in parapsychology.

9:30 "Effect of Agent's High and Low Aim" by Mr. D. Capuzzi (St. Joseph's College)

An experiment was designed to determine (1) whether high or low aim by the experimenter would affect the subject's ESP scoring level; and (2) whether ESP scores were related to personality traits as measured by the Minnesota Multiphasic Personality Inventory. Each of 35 subjects was given from one to four sets of ESP trials, each set consisting of two runs for telepathy and two for clairvoyance. There were no significant differences in the results based on the experimenter's aim. However, there were negative correlations between the ESP scores and four M.M.P.I. personality traits (psychasthenia, hypochondriasis, psychopathic deviate, and schizophrenia) which were significant at the 5 per cent level or better. Five of the other six traits were also negatively correlated, but not significantly.

10:00 "A Study of Moods and Attitude Variables as Related to Precognition" by Mrs. Winnifred Nielsen (University of Florida)[8]

The general objective of this research was to try to confirm previous evidence for the hypothesis that the mood of the subject is related to his precognition scores. To determine the subject's mood, the subject was asked to rate himself on his mental, emotional, and physical feelings; and from these a composite rating was derived which ranged on a five-point scale from "extremely pleasant" to "extremely unpleasant." Specific objectives of the research were concerned with whether the effect of mood on scores is different in different introversion-extraversion and attitude groups; whether extreme or moderate mood conditions are generally more favorable; and whether the pleasant-unpleasant dimension is an important factor. The over-all precognition scores were not significant. Further analyses planned on the basis of earlier findings support the hypothesis that the mood of the subject does affect his precognition scores. Extreme moods appear to have been more favorable for scoring by those subjects who believed in psi, an effect that was more pronounced for introverts than for extraverts. Among those who were not committed but were "open-minded" in their attitude towards the existence of ESP, the largest deviations were due to negative scores which occurred in extreme moods.

10:40 Symposium on Psychiatry and Parapsychology.

"Three Aspects of Psi in Psychotherapy" by Dr. Jan Ehrenwald (Roosevelt Hospital, N. Y.)

In the analysis of dreams recorded in the psychoanalytic setting, correspondences often show up between the items in the patient's dream and certain events in the therapist's own life or thoughts that are usual enough to indicate the presence of telepathy. The most obvious evidences of telepathy are found in the *uniqueness* of some of the items in the dream (such as dates, numbers, etc.) and in the presence of *tracer elements;* that is, a multiplicity of distinctive (though not necessarily unique) features in the dream which match real events in the therapist's life. Another criterion of the

[8] Presented by Miss Margaret Anderson.

presence of psi is that of *psychological significance.* In such cases, unless telepathy is assumed, the dream is unintelligible, its deeper meaning is obscured. This is a new principle, based on dynamic (rather than statistical or structural) considerations—one that is concerned with human motives and needs. The same is true when there is *doctrinal compliance;* that is, when the patient appears to be satisfying the therapist's own deep-seated needs by supplying, through his dream material, clinical evidence of the therapist's favorite hypotheses. The possibility of a direct telepathic impact of the therapist's personality upon the patient in effecting his cure —another aspect of psi in the therapeutic situation—is very difficult to prove or even to define.

"The Experimenter in Psi Investigation" by Dr. S. D. Kahn (U. S. Army Hospital, Fort Eustis, Va.)

In a field where there is frequent reference to the need for a repeatable experiment, one of the effects most often repeated is the decline and "disappearance" of successful experimenters. As Dr. Ehrenwald pointed out, in psychotherapy the unconscious needs of the analyst are often responded to by the patient and are of critical importance in eliciting his responses. It may be that the same applies in the subject-experimenter relationship. The fact that so many experimenters either fail to carry on the research they have started or to keep on getting good results if they do continue suggests that parapsychology could profit from a study of the effects that unconscious attitudes of the experimenter toward psi have upon the subject's performance.

"Telepathy in the Psychoanalytic Situation" by Dr. M. Ullman (New York University)

An analogy can be made between conditions for the occurrence of telepathic phenomena as they occur spontaneously and as they occur in the psychoanalytic setting. In both circumstances the conditions are characterized by the impossibility of normal communication and the existence of dire need or crisis. The difference is that in the spontaneous situation the conditions are real, objective, and external to the individual. In the psychoanalytic situation the distance is due

to psychological blockage and the need arises out of the neurotic structure of the patient's personality. In a small group of patients ("consistent telepathic dreamers") the ability to function at a telepathic level represents a major form of contact with the significant figures in their environment. The language and thought processes of these patients lose their effective connections with the external world, which results in desperate feelings of vulnerability and isolation. Telepathic contact occurs under these circumstances as a compensatory manoeuver.

GLOSSARY

A description of the basic experimental methods, of the findings, and of the statistical procedures for evaluating ESP and PK results may be found in *Parapsychology, Frontier Science of the Mind* by J. B. Rhine and J. G. Pratt (published by Charles C. Thomas, Springfield, Illinois, U.S.A., Blackwell Scientific Publications, Ltd., Oxford, England, and The Ryerson Press, Toronto, Canada).

AGENT: The "sender" in tests for telepathy, the person whose mental states are to be apprehended by the percipient. In GESP tests, the person who looks at the stimulus object.

BT (Basic Technique): The clairvoyance technique in which each card is laid aside by the experimenter as it is called by the subject. The check-up is made at the end of the run.

CALL: The ESP symbol selected by the subject in trying to identify a target.

CHANCE: The complex of undefined causal factors irrelevant to the purpose at hand.

Chance Expectation = Mean Chance Expectation: The most likely score if only chance variation is present.

Chance Average: Mean chance expectation expressed as an expected score, generally in terms of average per run.

CHI SQUARE: A sum of quantities each of which is a deviation squared divided by an expected value. Also a sum of the squares of CR's.

CLAIRVOYANCE: Extrasensory perception of objective events as distinguished from telepathic cognition of the mental activities of another person.

CR (Critical Ratio): A measure to determine whether or not the observed deviation is significantly greater than the expected random fluctuation about the average. The CR is obtained by dividing the observed deviation by the standard deviation. (The probability of a given CR may be obtained by consulting tables of the probability integral, such as Pearsons.)

CR_d (Critical Ratio of the Difference): The observed difference between the average scores of two samples of data divided by the standard deviation of the difference.

DEVIATION: The amount an observed number of hits or an average score varies from the mean chance expectation or chance average. A deviation may be total (for a series of runs) or average (per run).

DISPLACEMENT: ESP responses to targets other than those for which the calls were intended.

Backward Displacement: ESP responses to targets preceding the intended targets. Displacement to the targets one, two, three, etc. places preceding the intended target are designated as (—1), (—2), (—3), etc.

Forward Displacement: ESP responses to targets coming later than the intended targets. Displacement to the targets one, two, three, etc. places after the intended target are designated as (+1), (+2), (+3), etc.

DT (Down Through): The clairvoyance technique in which the cards are called down through the pack before any are removed or checked.

ESP (Extrasensory Perception): Awareness of or response to an external event or influence not apprehended by sensory means.

ESP CARDS: Cards, each bearing one of the following five symbols: star, circle, square, cross and waves (three parallel wavy lines).

ESP PACK: Twenty-five ESP cards.

Closed Pack: An ESP pack composed of five each of the five symbols.

Open Pack: An ESP pack made up of the ESP symbols selected in random order.

EXPECTATION: See ***Chance.***

EXTRACHANCE: Not due to chance alone.

FREE MATERIAL: The stimulus objects of experiments in which an unlimited or unspecified range of stimulus objects is employed (as contrasted with methods such as card-calling in which the subject knows that the stimulus object is one of a known range).

GESP (General Extrasensory Perception): A technique designed to test the occurrence of extrasensory perception, permitting either telepathy or clairvoyance or both to operate.

MCE (Mean Chance Expectation): see *Chance.*

P (Probability): A mathematical estimate of the expected relative frequency of a given event if chance alone were operative.

PARAPSYCHOLOGY: A division of psychology dealing with behavioral or personal effects that are demonstrably nonphysical (that is, which do not fall within the scope of known physical principles).

PERCIPIENT = SUBJECT: The person who makes the calls in an ESP test.

PK (Psychokinesis): The direct influence exerted on a physical system by a subject without any known intermediate physical energy or instrumentation.

PLACEMENT TEST: A PK technique in which the aim of the subject is to try to influence falling objects to come to rest in a designated part of the throwing surface.

PRECOGNITION: Cognition of a future event which could not be known through rational inference.

PREFERENTIAL MATCHING: A method of scoring responses to free material. A judge ranks the stimulus objects (usually pictures in sets of four) with respect to their similarity to, or association with, each response; and/or he ranks the responses with respect to their similarity to, or association with, each stimulus object.

PSI: A general term to identify personal factors or processes which are nonphysical in nature. It approximates the popular use of the word "psychic" and the technical one, "parapsychical."

PSI-MISSING: Exercise of psi ability in a way that avoids the target the subject is attempting to hit.

PSI PHENOMENA: Occurrences which result from the operation of psi. They include the phenomena of both ESP (including precognition) and PK.

PSYCHICAL RESEARCH: Older term used for parapsychology.

QD (Quarter Distribution): The distribution of hits in the record page (or in a logical subdivision thereof, such as the set or the half-set) as found in the four equal quarters formed by dividing the selected unit horizontally and vertically.

RANDOM ORDER: An order of events which displays no trends or regularities that would allow any inference regarding one event from one or more of the others in the series.

RUN: A group of trials, usually the successive calling of a deck of 25 ESP cards or symbols. In PK tests, 24 single die throws regardless of the number of dice thrown at the same time.

SCORE: The number of hits made in any given unit of trials.
Total Score: Total of scores made in a given number of runs.
Average Score: Total score divided by number of runs.

SD (Standard Deviation): Usually the theoretical root mean square of the deviations. It is obtained from the formula $\sqrt{npq}$ in which n is the number of single trials, p the probability of success per trial, and q the probability of failure.

SERIES: Several runs or experimental sessions that are grouped in accordance with a stated principle.

SESSION: A unit of an ESP or PK experiment comprising all the trials of one test occasion.

SET: A subdivision of the record page serving as a scoring unit for a consecutive group of trials, usually for the same target.

SIGNIFICANCE: A numerical result is significant when it equals or surpasses some criterion of degree of chance improbability. The criterion commonly used in parapsychology today is a probability value of .01 or less, or a deviation in either direction such that the CR is 2.58 or greater.

SINGLES TEST: A PK technique in which the aim of the subject is to try to influence dice to fall with a specified face up.

SPONTANEOUS PSI EXPERIENCE: Natural, unplanned occurrence of an event or experience that seems to involve parapsychical ability.

STIMULUS OBJECT: The ESP card or drawing or other object, some identifying characteristic of which is to be apprehended by the percipient.

STM (Screened Touch Matching): An ESP card test in which the subject indicates in each trial (by pointing to one of five key positions) what he thinks the top card is in the inverted pack held by the experimenter behind a screen. The card is then laid opposite that position.

SUBJECT: The person who is experimented upon. In ESP tests, most commonly the percipient (though also the agent in GESP and telepathy). In PK tests, any individual whose task it is to influence the objects thrown.

TARGET: In ESP (including precognition) the objective or mental events the subject is attempting to respond to. In PK tests, the faces of the die which the subject attempts to bring up in the act of throwing. In placement tests, the designated part of the throwing surface.

Target Card: The card which the percipient is attempting to identify or otherwise indicate a knowledge of.

Target Pack: The pack of cards the order of which the subject is attempting to identify.

Target Face: The face on the die which the subject tries to turn up by PK.

TELEPATHY: Extrasensory perception of the mental activities of another person (as distinguished from clairvoyant cognition of objective events).

TRIAL: In ESP tests, a single attempt to identify a stimulus object. In PK tests, a single throw of the dice or other objects thrown.